



**LOURENÇO CRUZ SIMÕES DURÃO**

Licenciado em Engenharia de Energias e Ambiente

# **DETEÇÃO DE AVARIAS EM BATERIAS DE SEGUIDORES FOTOVOLTAICOS POR ANÁLISE DE SÉRIES TEMPORAIS REAIS**

MESTRADO EM ENGENHARIA DE ENERGIAS RENOVÁVEIS

Universidade NOVA de Lisboa  
outubro, 2023

# DETEÇÃO DE AVARIAS EM BATERIAS DE SEGUIDORES FOTOVOLTAICOS POR ANÁLISE DE SÉRIES TEMPORAIS REAIS

LOURENÇO CRUZ SIMÕES DURÃO

Licenciado em Engenharia de Energias e Ambiente

**Orientador:** João Murta Pina

*Professor Auxiliar, Universidade NOVA de Lisboa*

**Coorientador:** Thiago Stefanutto

*Solar Specialist Asset Performance, GALP*

## Júri

**Presidente:** Daniel Cardoso Vaz  
*FCT-NOVA*

**Arguente:** Rui Miguel Amaral Lopes  
*FCT-NOVA*

**Vogal:** João Miguel Murta Pina  
*FCT-NOVA*

## **Deteção de avarias em baterias de seguidores fotovoltaicos por análise de séries temporais reais**

Copyright © Lourenço Cruz Simões Durão, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

## AGRADECIMENTOS

Gostaria de expressar meus sinceros agradecimentos às seguintes pessoas e instituições que desempenharam papéis fundamentais no desenvolvimento e conclusão desta tese:

À GALP, em especial, pelo acolhimento e pela oportunidade de realizar um estágio nas suas instalações. Sou imensamente grato pela autorização concedida para explorar os campos fotovoltaicos da empresa neste trabalho de mestrado.

À Samsung, sou imensamente grato pelo curso de inteligência artificial que me proporcionaram. Sem esse curso, eu não teria adquirido o conhecimento e as habilidades necessárias para desenvolver a minha tese, que está intimamente relacionada com esse campo.

Aos meus pais, cujo amor e incentivo me sustentaram durante toda a jornada acadêmica. Suas palavras de encorajamento e crença em mim foram a motivação para superar os desafios e me esforçar ao máximo.

À minha Tia Catarina e a minha Madrinha, por sua orientação e apoio contínuos. Suas valiosas sugestões e feedbacks enriqueceram meu trabalho, ajudando-me a aprimorar minhas ideias e a estrutura da tese.

Aos meus amigos, cuja amizade e apoio incondicionais foram essenciais durante todo o período de pesquisa. Seus estímulos, conversas inspiradoras e momentos de descontração trouxeram equilíbrio à minha vida acadêmica.

Expresso minha profunda gratidão a todos vocês. Seu apoio e contribuições desempenharam um papel fundamental no meu sucesso. Sou extremamente privilegiado por tê-los ao meu lado nesta jornada e serei eternamente grato por tudo o que fizeram por mim.

”

«*No man was ever wise by chance.*»

— Seneca

## ABSTRACT

The unpredictable malfunctions of batteries in photovoltaic trackers pose significant challenges in terms of energy efficiency and operational costs. This thesis addresses the need for an effective solution to determine whether battery alarms are anomalies or actual failures, enabling swift replacements and substantial energy savings. Given the scarcity of studies in this area, the proposed research aims to develop a program utilizing artificial intelligence algorithms to identify patterns characteristic of end-of-life batteries. By leveraging artificial intelligence techniques, both economic and energy gains can be achieved through timely and accurate battery replacement, thereby reducing unpredictability in photovoltaic tracker systems.

**Keywords:** Artificial Intelligence, Battery, Long-Short Term Memory, Predictive Maintenance, Tracker Control Unit

## RESUMO

As avarias imprevisíveis das baterias dos seguidores fotovoltaicos colocam desafios significativos em termos de eficiência energética e de custos operacionais. Esta tese aborda a necessidade de uma solução eficaz para determinar se os alarmes das baterias são anomalias ou falhas reais, permitindo substituições rápidas e poupanças substanciais de energia. Dada a escassez de estudos nesta área, a investigação proposta visa desenvolver um programa que utiliza algoritmos de inteligência artificial para identificar padrões característicos de baterias em fim de vida. Ao utilizar técnicas de inteligência artificial, é possível obter ganhos económicos e energéticos através da substituição atempada e precisa das baterias, reduzindo assim a imprevisibilidade dos sistemas de seguidores fotovoltaicos.

**Palavras-chave:** Bateria, Inteligencia Artificial, *Long-Short Term Memory*, Manutenção Preditiva, Seguidores Fotovoltaicos,

# ÍNDICE

|                                                                                                                        |             |
|------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>Índice de Figuras</b>                                                                                               | <b>viii</b> |
| <b>Glossário</b>                                                                                                       | <b>x</b>    |
| <b>Siglas</b>                                                                                                          | <b>xii</b>  |
| <b>Símbolos</b>                                                                                                        | <b>xiv</b>  |
| <b>1 Introdução</b>                                                                                                    | <b>1</b>    |
| 1.1 Enquadramento . . . . .                                                                                            | 1           |
| 1.2 Objetivos . . . . .                                                                                                | 2           |
| 1.3 Contributos originais . . . . .                                                                                    | 3           |
| 1.4 Descrição da organização da Tese . . . . .                                                                         | 3           |
| <b>2 Estado da Arte</b>                                                                                                | <b>5</b>    |
| 2.1 Introdução . . . . .                                                                                               | 5           |
| 2.2 Conceitos Gerais e Tipos de Manutenção . . . . .                                                                   | 5           |
| 2.2.1 Manutenção Preventiva . . . . .                                                                                  | 7           |
| 2.2.2 Manutenção Preditiva . . . . .                                                                                   | 9           |
| 2.2.3 Manutenção Corretiva . . . . .                                                                                   | 10          |
| 2.2.4 Manutenção Baseada na Condição . . . . .                                                                         | 10          |
| 2.3 Manutenção Preditiva com Inteligencia Artificial . . . . .                                                         | 11          |
| 2.3.1 O Conceito de Inteligência Artificial e Machine Learning e a sua<br>Relação com a Manutenção Preditiva . . . . . | 11          |
| 2.3.2 Benefícios do Uso de IA para a Manutenção Preditiva . . . . .                                                    | 12          |
| 2.3.3 Tipos de Modelos Usados de IA para a Manutenção Preditiva . . . . .                                              | 12          |
| 2.3.4 Limitações à Implementação da IA na Manutenção Preditiva . . . . .                                               | 22          |
| 2.3.5 Aplicação da Inteligência Artificial na Manutenção Preditiva: Um<br>Estudo de Caso . . . . .                     | 23          |
| 2.4 Manutenção Preditiva em Instalações Fotovoltaicas . . . . .                                                        | 25          |



|          |                                                                   |           |
|----------|-------------------------------------------------------------------|-----------|
| 2.5      | Inteligencia Artificial em Baterias . . . . .                     | 28        |
| <b>3</b> | <b>Metodologia</b>                                                | <b>31</b> |
| <b>4</b> | <b>Caso de Estudo</b>                                             | <b>36</b> |
| <b>5</b> | <b>Análise dos Resultados</b>                                     | <b>38</b> |
| 5.1      | Correlações . . . . .                                             | 38        |
| 5.2      | Visualização dos dados e criação de uma <i>Baseline</i> . . . . . | 40        |
| 5.3      | Resultados do Modelo LSTM Multivariável . . . . .                 | 42        |
| 5.4      | Resultados do Modelo LSTM Univariável . . . . .                   | 48        |
| 5.5      | Simulação Económica . . . . .                                     | 54        |
| <b>6</b> | <b>Conclusão</b>                                                  | <b>55</b> |
| 6.1      | Conclusão . . . . .                                               | 55        |
| 6.2      | Trabalhos Futuros . . . . .                                       | 56        |
|          | <b>Bibliografia</b>                                               | <b>58</b> |

## ÍNDICE DE FIGURAS

|      |                                                                                                                                     |    |
|------|-------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1  | Imagem do TCU (1) e do motor (2). . . . .                                                                                           | 1  |
| 2.1  | Número de documentos relativos a certas estratégias de manutenção por ano de 2000 a 2018; Retirada de Basri et al. (2017) . . . . . | 7  |
| 2.2  | Estrutura Árvore de Decisão; Retirada de Rokach e Maimon (2005). . . . .                                                            | 13 |
| 2.3  | Estrutura da Floresta Aleatória; Retirada de Ouadah et al. (2022). . . . .                                                          | 15 |
| 2.4  | Máquina de Vetores de Suporte; Retirada de Ouadah et al. (2022). . . . .                                                            | 15 |
| 2.5  | Exemplo de Rede Neuronal Artificial; Retirada de Nielsen (2015) . . . . .                                                           | 17 |
| 2.6  | Esquema de Funcionamento e Classificação dos Gates; Retirada de Zbiciak e Markiewicz (2023) . . . . .                               | 19 |
| 2.7  | Ilustração dos Processos Matemáticos de uma Célula LSTM. . . . .                                                                    | 20 |
| 2.8  | Uma arquitetura CNN simples, composta por cinco camadas; Retirada de O'Shea e Nash (2015) . . . . .                                 | 22 |
| 2.9  | Tipos de Modelos Utilizados; Retirada de Betti et al. (2019) . . . . .                                                              | 27 |
| 2.10 | Métricas para Avaliação de Diferentes Métodos; Retirada de J. C. Chen et al. (2021) . . . . .                                       | 30 |
| 2.11 | a) Previsão de Vida Remanescente Útil para B0018 com SP = 90% b) Erro de Previsão; Retirada de J. C. Chen et al. (2021) . . . . .   | 30 |
| 4.1  | Parque "Alcazar"190MWp. . . . .                                                                                                     | 36 |
| 4.2  | Descrição das características principais dos parques fotovoltaicos . . . . .                                                        | 36 |
| 4.3  | Transformador e Inversor de "Alcazar". . . . .                                                                                      | 37 |
| 4.4  | Subestação de "Alcazar". . . . .                                                                                                    | 37 |
| 5.1  | Correlação das variáveis com maior influência no nível de bateria. . . . .                                                          | 39 |
| 5.2  | Gráficos das Variáveis Usadas para a Criação do Modelo LSTM Multivariada. . . . .                                                   | 40 |
| 5.3  | Regressão Linear - Previsão vs Valores reais. . . . .                                                                               | 41 |
| 5.4  | Tabela de métricas obtidas pela regressão linear. . . . .                                                                           | 41 |
| 5.5  | LSTM univariável - Previsão vs Valores reais. . . . .                                                                               | 42 |
| 5.6  | LSTM univariável - Previsão vs Valores reais. . . . .                                                                               | 42 |

|      |                                                                                                                 |    |
|------|-----------------------------------------------------------------------------------------------------------------|----|
| 5.7  | Previsão vs Real Bateria Funcional 1. . . . .                                                                   | 43 |
| 5.8  | Previsão vs Real Bateria Funcional 2. . . . .                                                                   | 43 |
| 5.9  | Previsão vs Real Bateria Funcional 3. . . . .                                                                   | 43 |
| 5.10 | Previsão vs Real Bateria Funcional 4. . . . .                                                                   | 44 |
| 5.11 | Previsão vs Real Bateria Funcional 5. . . . .                                                                   | 44 |
| 5.12 | Previsão vs Real Bateria Funcional 6. . . . .                                                                   | 44 |
| 5.13 | Previsão vs Real Bateria Funcional 7. . . . .                                                                   | 45 |
| 5.14 | Previsão vs Real Bateria Funcional 8. . . . .                                                                   | 45 |
| 5.15 | Previsão vs Real Bateria Funcional 9. . . . .                                                                   | 45 |
| 5.16 | Previsão vs Real Bateria Funcional 10. . . . .                                                                  | 46 |
| 5.17 | Previsão vs Real Bateria Funcional 11. . . . .                                                                  | 46 |
| 5.18 | Previsão vs Real Bateria Avariada A. . . . .                                                                    | 46 |
| 5.19 | Previsão vs Real Bateria Avariada B. . . . .                                                                    | 47 |
| 5.20 | Previsão vs Real Bateria Avariada C. . . . .                                                                    | 47 |
| 5.21 | Métricas de Previsão VS Real do Modelo LSTM Multivariável. . . . .                                              | 47 |
| 5.22 | Previsão vs Real Bateria Funcional 1. . . . .                                                                   | 48 |
| 5.23 | Previsão vs Real Bateria Funcional 2. . . . .                                                                   | 49 |
| 5.24 | Previsão vs Real Bateria Funcional 3. . . . .                                                                   | 49 |
| 5.25 | Previsão vs Real Bateria Funcional 4. . . . .                                                                   | 49 |
| 5.26 | Previsão vs Real Bateria Funcional 5. . . . .                                                                   | 50 |
| 5.27 | Previsão vs Real Bateria Funcional 6. . . . .                                                                   | 50 |
| 5.28 | Previsão vs Real Bateria Funcional 7. . . . .                                                                   | 50 |
| 5.29 | Previsão vs Real Bateria Funcional 8. . . . .                                                                   | 51 |
| 5.30 | Previsão vs Real Bateria Funcional 9. . . . .                                                                   | 51 |
| 5.31 | Previsão vs Real Bateria Funcional 10. . . . .                                                                  | 51 |
| 5.32 | Previsão vs Real Bateria Funcional 11. . . . .                                                                  | 52 |
| 5.33 | Previsão vs Real Bateria Avariada A. . . . .                                                                    | 52 |
| 5.34 | Previsão vs Real Bateria Avariada B. . . . .                                                                    | 52 |
| 5.35 | Previsão vs Real Bateria Avariada C. . . . .                                                                    | 53 |
| 5.36 | Métricas de Previsão VS Real do Modelo LSTM Univariável. . . . .                                                | 53 |
| 5.37 | Simulação Económica das perdas energéticas por falhas dos seguidores foto-<br>voltaicos no “Alcazar 1”. . . . . | 54 |

## GLOSSÁRIO

|                             |                                                                                                                                                                                                                               |
|-----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>bias</b>                 | Refere-se aos parâmetros adicionais que são introduzidos nas células de memória para ajudar na capacidade de aprendizagem da rede. (p. 19)                                                                                    |
| <b>deep learning</b>        | É um subcampo da inteligência artificial e de <i>Machine Learning</i> que se concentra no treino de redes neurais artificiais para realizar tarefas complexas de forma autônoma. (p. 23)                                      |
| <b>função Kernel</b>        | Função matemática que permite mapear os dados de um espaço de características original para um espaço de características de maior dimensão. (p. 15)                                                                           |
| <b>hiperparâmetros</b>      | São parâmetros externos ao modelo de aprendizagem que controlam o comportamento e a configuração do algoritmo de treino. (p. 16)                                                                                              |
| <b>hiperplano</b>           | Refere-se a uma superfície de decisão que separa os pontos de dados em diferentes classes. É uma representação geométrica que divide o espaço de características em duas regiões, com uma classe em cada região. (p. 15)      |
| <b>otimização Bayesiana</b> | É uma técnica usada para encontrar a configuração ideal de hiperparâmetros de um modelo de <i>machine learning</i> . (p. 28)                                                                                                  |
| <b>overfitting</b>          | Refere-se a um problema de ajuste. O overfitting ocorre quando um modelo se torna muito específico para os dados de treino utilizados, capturando até mesmo o ruído e as variações aleatórias presentes nesses dados. (p. 14) |
| <b>pesos</b>                | Referem-se aos parâmetros que são ajustados durante o processo de treino da rede. (p. 17)                                                                                                                                     |

**string**

Um conjunto de dispositivos eletrônicos interconectados em série.  
(*p.* [2](#))

## SIGLAS

|              |                                                         |
|--------------|---------------------------------------------------------|
| <b>ANN</b>   | <i>Artificial Neural Network (p. 16)</i>                |
| <b>AVAC</b>  | <i>Aquecimento, Ventilação e Ar Condicionado (p. 9)</i> |
| <b>CNN</b>   | <i>Convolucional Neural Network (p. 16)</i>             |
| <b>FNN</b>   | <i>Feedforward Neural Network (p. 16)</i>               |
| <b>FV</b>    | <i>Fotovoltaicos (p. 1)</i>                             |
| <b>GPU</b>   | <i>Graphics Processing Unit (p. 35)</i>                 |
| <b>IA</b>    | <i>Inteligência Artificial (p. 5)</i>                   |
| <b>LSTM</b>  | <i>Long Short Term Memory (p. 17)</i>                   |
| <b>MAE</b>   | <i>Mean Absolute Error (p. 35)</i>                      |
| <b>MAPE</b>  | <i>Mean Absolute Percentage Error (p. 35)</i>           |
| <b>MBC</b>   | <i>Manutenção com Base na Condição (p. 10)</i>          |
| <b>ML</b>    | <i>Machine Learning (p. 5)</i>                          |
| <b>MSE</b>   | <i>Mean Squared Error (p. 35)</i>                       |
| <b>PMS</b>   | <i>Plant Monitoring System (pp. 31, 35)</i>             |
| <b>RMSE</b>  | <i>Root Mean Squared Error (p. 35)</i>                  |
| <b>RNN</b>   | <i>Recurrent Neural Network (p. 16)</i>                 |
| <b>RUL</b>   | <i>Remaining Useful Life (p. 28)</i>                    |
| <b>SCADA</b> | <i>Supervisory Control and Data Acquisition (p. 27)</i> |
| <b>SDM</b>   | <i>Supervision Diagnosis Model (p. 27)</i>              |
| <b>SOC</b>   | <i>State Of Change (p. 3)</i>                           |

|              |                                                                 |
|--------------|-----------------------------------------------------------------|
| <b>SP</b>    | <i>Starting Positions (p. <a href="#">28</a>)</i>               |
| <b>STFPM</b> | <i>Short Term Fault Predction Model (p. <a href="#">27</a>)</i> |
| <b>SVM</b>   | <i>Support Vector Machine (p. <a href="#">15</a>)</i>           |

## SÍMBOLOS

|             |                                                                                                   |
|-------------|---------------------------------------------------------------------------------------------------|
| $\dot{C}$   | Percentagem de importância para o passo de tempo atual (pp. 20, 21)                               |
| $C_{(t)}$   | Estado da célula Atual (p. 21)                                                                    |
| $C_{(t-1)}$ | Estado anterior da célula (p. 21)                                                                 |
| $E^f$       | Enviesamento de Vetor de Entrada Atual na entrada f (pp. 20, 21)                                  |
| $E^g$       | Enviesamento de Vetor de Entrada Atual na entrada g (pp. 20, 21)                                  |
| $E^h$       | Enviesamento de Vetor de Entrada Atual na entrada h (p. 21)                                       |
| $E^t$       | Enviesamento de Vetor de Entrada Atual na entrada t (p. 21)                                       |
| $F_{(t)}$   | Percentagem de Estado anterior da célula deve ser retido para o passo de tempo atual (pp. 20, 21) |
| $h_{(t)}$   | Estado Oculto Atual (p. 21)                                                                       |
| $h_{(t-1)}$ | Estado Oculto Anterior (pp. 20, 21)                                                               |
| $I$         | Percentagem de vetor candidato que é retido para o estado anterior da célula (p. 21)              |
| $L$         | Memória potencial de longo prazo (p. 21)                                                          |
| $N$         | Memória potencial de curto prazo (p. 21)                                                          |
| $O$         | Percentagem do estado oculto anterior que deve ser produzido na etapa de tempo atual (p. 21)      |
| $U^f$       | Pesos de Vetor de Entrada Atual na entrada f (pp. 20, 21)                                         |
| $U^g$       | Pesos de Vetor de Entrada Atual na entrada g (pp. 20, 21)                                         |
| $U^h$       | Pesos de Vetor de Entrada Atual na entrada h (p. 21)                                              |
| $U^t$       | Pesos de Vetor de Entrada Atual na entrada t (p. 21)                                              |



|       |                                                           |
|-------|-----------------------------------------------------------|
| $W^f$ | Pesos de Estado Oculto Anterior na entrada f (pp. 20, 21) |
| $W^g$ | Pesos de Estado Oculto Anterior na entrada g (pp. 20, 21) |
| $W^h$ | Pesos de Estado Oculto Anterior na entrada h (p. 21)      |
| $W^t$ | Pesos de Estado Oculto Anterior na entrada t (p. 21)      |
| $X_b$ | Valor de Entrada na Função sigmoide b (p. 21)             |
| $X_f$ | Valor de Entrada na Função sigmoide f (pp. 20, 21)        |
| $X_g$ | Valor de Entrada na Função Tanh (pp. 20, 21)              |
| $X_h$ | Valor de Entrada na Função sigmoide h (p. 21)             |
| $x_t$ | Vetor de entrada atual (pp. 20, 21)                       |

# INTRODUÇÃO

Este capítulo foca-se na descrição do problema que se manifesta nos parques fotovoltaicos equipados com seguidores fotovoltaicos. Além disso, delineará o objetivo da presente tese, que consiste em solucionar esse problema;

## 1.1 Enquadramento

Para o desenvolvimento desta tese, foi analisado o problema de manutenção das baterias dos seguidores **Fotovoltaicos (FV)**. Caracterizadas pela sua imprevisibilidade e recorrente necessidade de manutenção, é, portanto, o objetivo desta tese possibilitar a previsão de falhas de funcionamento e, consequentemente, otimizar o processo de manutenção.

As centrais fotovoltaicas utilizam seguidores, denominados Tracker Control Unit (TCU), por forma a maximizar a energia captada do sol. Cada TCU está conectado a um motor responsável por girar e ajustar a estrutura que suporta os painéis fotovoltaicos como podemos ver pela Figura 1.1, posicionando-os de forma perpendicular à radiação solar. Os TCUs são compostos por uma bateria e um comunicador sem-fios, e encontram-se ligados a um Network Control Unit (NCU), responsável pela supervisão e coordenação de conjuntos de TCUs.

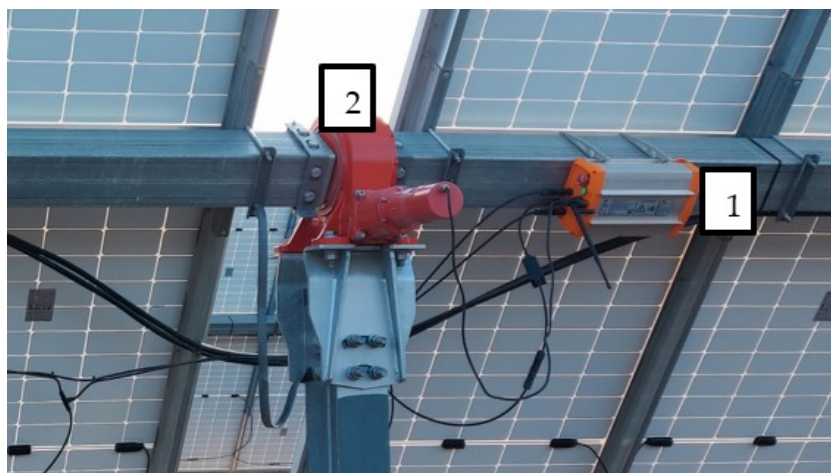


Figura 1.1: Imagem do TCU (1) e do motor (2).

Embora o TCU possa identificar algumas condições de alarme, tais como bloqueios e baterias fracas, não é capaz de detetar riscos de vento ou neve. Assim os NCU's utilizam indicadores de uma estação meteorológica, incluindo temperatura, humidade, velocidade e direção do vento, radiação solar direta e indireta, e precipitação, permitindo coordenar os TCUs. Os NCU's monitorizam e gerem os alarmes de vento e neve, enviando pedidos de posição segura a todos os TCUs afetados. Cada NCU pode ter até 224 TCUs conectados. Cada TCU possui um painel fotovoltaico independente para recarregar a sua bateria pois a dimensão dos módulos independentes é diferente do resto dos módulos.

Estes painéis fotovoltaicos seguem o movimento do sol ao longo do dia, exceto quando é necessário protegê-los de ventos fortes. Importante salientar que o painel independente está fixado na mesma estrutura dos painéis fotovoltaicos, embora a sua ligação seja independente da [string](#) dos módulos fotovoltaicos.

### 1.2 Objetivos

A velocidade na substituição das baterias dos TCUs é crucial para uma melhor produtividade das centrais FV. Por forma a otimizar a utilização dos painéis e recolher o máximo de energia solar possível, torna-se indispensável uma rápida substituição das baterias, para que os motores possam estar em pleno funcionamento. Em virtude do impacto que as baterias dos seguidores (TCUs) têm na produtividade das centrais FV, esta tese tem como objetivo a criação de um método para diminuir o tempo de substituição das avarias, assim como reduzir a subjetividade nos parâmetros de determinação de uma avaria de bateria.

Detetar o comportamento de uma bateria no fim da sua vida útil, permitir-nos-á entender melhor como e quando devemos proceder a uma reparação e/ou substituição das baterias. Utilizando uma base de dados e modelos de inteligência artificial, determinaremos o comportamento.

Nos parques fotovoltaicos, são frequentes as avarias dos TCUs associadas às baterias, resultando numa produção deficiente dos módulos fotovoltaicos e, nalguns casos, numa produção inferior à de um painel sem TCU.

Os principais desafios destas avarias são a falta de previsibilidade e o protocolo de manutenção, que exige que no momento que exista um alarme de falha de bateria, que significa que o nível de bateria é inferior a 25%, a reinicialização da bateria e uma espera de dois dias que é considerado o tempo adequado para verificar se a bateria retorna ao seu funcionamento normal. Nalguns casos, a bateria volta a funcionar normalmente, enquanto noutras ocasiões, funciona bem por alguns dias e ainda noutras não funciona após os dois dias, levando à substituição do TCU. No entanto, a falha no funcionamento da bateria é um problema subjetivo, uma vez que é comum as baterias apresentarem níveis de bateria diferentes umas das outras.

O objetivo deste trabalho é eliminar a subjetividade destes problemas por meio da utilização da inteligência artificial para determinar se é necessária a substituição do TCU. A abordagem utilizada neste estudo consiste na criação de uma rede neural artificial que

emprega um modelo LSTM para verificar o funcionamento da bateria antes do surgimento dos primeiros problemas, utilizando vários parâmetros e observando as variações no nível de bateria (*State Of Charge* (SOC)). Dessa forma, será possível verificar se o comportamento observado 48 horas antes do primeiro alarme e nas 24 horas seguintes se é semelhante ao comportamento de uma bateria no fim de vida. Isso permitirá que os técnicos em campo tenham uma base mais objetiva e um valor de comparação para tomar decisões sobre a substituição da bateria.

### 1.3 Contributos originais

Os contributos originais desta pesquisa residem na aplicação inovadora de uma metodologia baseada em modelos LSTM no contexto dos seguidores fotovoltaicos. Com esta abordagem, foi possível identificar de forma eficaz o fim de vida da bateria desses dispositivos, fornecendo uma solução relevante para um problema desafiante. Esta aplicação pioneira dos modelos LSTM ao tema dos seguidores fotovoltaicos representa um avanço significativo no campo, ampliando o conhecimento e abrindo possibilidades para futuras pesquisas e aplicações práticas.

### 1.4 Descrição da organização da Tese

#### Capítulo 2: Estado da Arte

Neste capítulo, é feita uma revisão detalhada da literatura relacionada com o tema da pesquisa. Introduz os conceitos gerais e os diferentes tipos de manutenção, como a preventiva, preditiva, corretiva e baseada na condição. Aborda a aplicação da manutenção preditiva, utilizando inteligência artificial, explorando o conceito de inteligência artificial e *machine learning* e a sua relação com a manutenção preditiva. Também discute os benefícios, tipos de modelos e limitações do uso de IA e ML para a manutenção preditiva. Além disso, explora a manutenção preditiva em instalações fotovoltaicas e a aplicação de inteligência artificial em baterias.

#### Capítulo 3: Metodologia

Este capítulo descreve a metodologia utilizada no trabalho, detalha os procedimentos e abordagens adotados para a recolha e análise de dados, assim como a seleção dos modelos e técnicas utilizados.

#### Capítulo 4: Caso de Estudo

Neste capítulo, é apresentado o caso de estudo específico analisado na pesquisa. Também destaca as características e particularidades do caso de estudo em relação ao tema da tese.

#### Capítulo 5: Análise de Resultados

Este capítulo apresenta os resultados do trabalho. Começa por descrever as correlações identificadas entre as variáveis estudadas. Em seguida, aborda a visualização dos dados e a criação de uma *baseline* como referência. Também discute os resultados obtidos com a aplicação do modelo LSTM multivariável e do modelo LSTM univariável. Além disso, inclui uma simulação económica relacionada com o tema.

### **Capítulo 6: Conclusão**

No capítulo final, são apresentadas as conclusões do estudo. Resumem-se os principais resultados alcançados, discute-se a relevância desses resultados em relação aos objetivos da pesquisa e destacam-se possíveis direções futuras para pesquisas relacionadas com o tema

## ESTADO DA ARTE

### 2.1 Introdução

Neste capítulo, serão apresentados e analisados os conceitos gerais de vários tipos de manutenção, como a manutenção preventiva, corretiva, preditiva de forma genérica e baseada na condição. Além disso, serão abordados os conceitos relacionados, tais como, como manutenção preditiva utilizando inteligência artificial.

Será discutido o conceito de *Inteligência Artificial (IA)* e *Machine Learning (ML)*, bem como a sua relação com a manutenção preditiva. Serão destacados os benefícios do uso de IA para a manutenção preditiva, assim como os tipos de modelos utilizados nesse contexto. Também serão abordadas as limitações à implementação da IA na manutenção preditiva.

Adicionalmente, serão exploradas a manutenção preditiva em instalações fotovoltaicas e a aplicação da inteligência artificial em baterias. Será discutida a relevância dessas aplicações específicas para aprimorar a eficiência e o desempenho dos sistemas.

Ao longo deste capítulo, serão examinados de forma abrangente estes tópicos fundamentais, fornecendo uma compreensão aprofundada da manutenção e da aplicação da inteligência artificial na manutenção preditiva.

### 2.2 Conceitos Gerais e Tipos de Manutenção

A manutenção é um aspecto vital de qualquer indústria, pois garante o bom funcionamento e a longevidade do equipamento e da maquinaria. A importância da manutenção na indústria não pode ser subestimada, pois não só garante o bom funcionamento dos processos como também ajuda a minimizar os tempos de paragem, a reduzir custos e a melhorar a segurança.

Os custos de manutenção são uma grande preocupação para os gestores das empresas. É importante lembrar que os custos associados à manutenção podem chegar a ser entre 15-60% dos custos totais de seus produtos. Em setores como a agricultura, os custos da manutenção em relação ao custo total dos bens produzidos são de 15%, mas em setores

como o metalúrgico rondam os 60% do custo do produto («Contents», 2002).

As empresas produtoras de bens, que gerem atualmente um negócio, enfrentam exigências cada vez maiores na área da fiabilidade das infraestruturas de produção e na garantia de operações indefetíveis, que são afetadas por tempos de paragem não programados. Não é fácil manter a fiabilidade de produção devido a eventos imprevistos que possam ocorrer durante o tempo de funcionamento do equipamento. A grande dificuldade está relacionada com um elevado nível de incerteza em relação às paragens de produção e à antecipação de tais ameaças com a devida reserva de tempo (Gackowiec, 2019).

Perturbações no bom funcionamento das instalações, com elevada probabilidade, levam a graves perdas de produtividade, declínios de qualidade, e também a deteriorações notáveis no desempenho global. A prevenção dos mencionados defeitos e falhas indesejáveis das máquinas, bem como a redução dos efeitos adversos por elas causados, são possíveis através da garantia de ações de manutenção corretas no sistema de gestão da produção. O desenvolvimento na indústria contemporânea e as mudanças na abordagem dos processos de fabrico, para além do progresso da mecanização, informatização e dispositivos tecnológicos altamente avançados, criam desafios para os engenheiros de manutenção assumirem. A manutenção é um termo relacionado com todas as atividades e procedimentos, planeados e não planeados, que são empreendidos para assegurar a acessibilidade constante do equipamento operacional numa fábrica de produção (Gackowiec, 2019).

A manutenção é baseada no planeamento e utiliza as 3 facetas que foram anteriormente faladas, este é categorizado de três formas de acordo com (Basri et al., 2017):

- Planeamento Baseado em Custos;
- Planeamento Baseado no Tempo;
- Planeamento Baseado em Falhas;

O planeamento baseado em **custos** analisa os custos e benefícios de capital para as organizações que utilizam manutenção preventiva, bem como as receitas que estas ajudam a gerar (Basri et al., 2017).

O planeamento baseado no **tempo** envolve a medição da análise do tempo em termos da atribuição e informação de um determinado período ou duração de uma operação. A base temporal é crucial no planeamento, uma vez que serve uma análise que indica a quantidade de tempo gasto em projetos, geralmente em referência a processos periódicos (Basri et al., 2017).

O planeamento baseado em **falhas**, envolve análises que têm em conta informações sobre a deterioração do sistema, ou dos componentes. No planeamento baseado em falhas, com o objetivo de antecipar as características de um mecanismo de falha, são realizadas análises críticas envolvendo os objetivos da análise e avaliações de fatores de falha, as quais

são amplamente quantificadas. Isto deverá assegurar que as manutenções preventivas sejam conduzidas corretamente, reduzindo assim as falhas (Basri et al., 2017).

Existem vários tipos de manutenção, mas com base em Gackowiec (2019) pode-se ver na Figura 2.1 quais são as estratégias que aparecem com maior frequência em trabalhos académicos: manutenção preventiva, manutenção preditiva, manutenção corretiva, manutenção com base na condição e manutenção pró-ativa. As técnicas para as políticas de manutenção que serão o foco do estudo bibliográfico, estarão de acordo com Çinar et al. (2020) e são: manutenção preventiva, manutenção preditiva, manutenção corretiva, manutenção com base na condição. Decidiu-se não efetuar uma análise detalhada da manutenção proativa devido à sua pouca relevância.

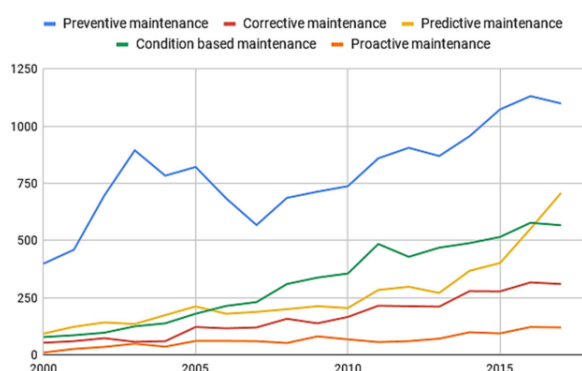


Figura 2.1: Número de documentos relativos a certas estratégias de manutenção por ano de 2000 a 2018; Retirada de Basri et al. (2017)

### 2.2.1 Manutenção Preventiva

A manutenção preventiva é um tipo de manutenção que é realizada em equipamentos ou máquinas, a fim de evitar potenciais falhas ou avarias. Isto é conseguido através da realização de inspeções regulares, limpeza e testes do equipamento e da substituição de peças desgastadas ou danificadas antes de causarem uma avaria (Basri et al., 2017). O objetivo da manutenção preventiva é prolongar a vida útil do equipamento, minimizar o tempo de paragem e reduzir a necessidade de reparações mais dispendiosas.

Com o objetivo de implementar eficazmente um programa de manutenção preventiva, é importante estabelecer um calendário para inspeções e manutenção regulares (Basri et al., 2017). Este programa deve basear-se nas recomendações do fabricante e nas necessidades específicas do equipamento ou maquinaria em questão. Este calendário é estabelecido através de dados históricos e através de estatística. Além disso, é importante ter um sistema de rastreio e documentação da manutenção que é realizada. Isto ajudará a assegurar que toda a manutenção necessária é executada dentro do prazo e que quaisquer potenciais problemas são identificados e corrigidos de forma atempada (Basri et al., 2017).

A manutenção preventiva é realizada sob duas perspetivas, que são conhecidas como gestão e operacional (Simões et al., 2011). A perspetiva de gestão refere-se ao apoio à tomada de decisões que facilita a análise dos dados. Os contributos para a perspetiva



de gestão incluem a determinação de objetivos de manutenção preventiva, planeamento de ações de manutenção e métodos envolvidos na resolução de qualquer problema que ocorra em relação a manutenção preventiva, bem como o desempenho dos sistemas. A perspectiva de gestão é também conhecida como um processo externo, uma vez que as suas decisões se baseiam na história e análise anteriores à execução de ações de manutenção preventiva (Söderholm et al., 2007).

A perspectiva operacional refere-se à execução de ações de manutenção, a fim de sustentar a capacidade de um sistema para desempenhar as funções pretendidas. Esta perspectiva é um processo interno que consiste em aspetos técnicos através dos quais as manutenções preventivas são executadas com base em entradas para o processo externo (Söderholm et al., 2007).

O conceito de planeamento manutenção preventiva que é discutido, consiste em três facetas conforme Young (2003):

- Objetivo(s) ou a finalidade de realizar o planeamento de manutenção preventiva;
- Descrições do estado do sistema em termos da sua importância e das suas funções;
- Métodos divididos em várias classificações que ajudam a determinar as melhores soluções para as questões destacadas (Basri et al., 2017);

Como parte do conceito de planeamento, "**finalidade**" é descrita como o objetivo ou meta pretendida através da realização do planeamento da manutenção preventiva. A razão para estabelecer objetivos significativos antes da execução do planeamento de manutenção preventiva é reduzir e concentrar o estudo e orientar a recolha de informação pertinente (Basri et al., 2017).

O conceito de planeamento manutenção preventiva engloba a **descrição do estado**, função e importância de um sistema. Um sistema está previsto como um conjunto de componentes interligados dispostos para levar a cabo um processo. De uma perspectiva operacional, o sistema é definido em termos do seu estado. O objetivo de considerar o estado de um sistema é representá-lo como operando normalmente, como operando em modo de avaria, ou como tendo falhado. Assim, a tomada de decisão para o planeamento e manutenção preventiva baseia-se no estado do sistema, que procura basear uma análise sobre o estado do sistema no que diz respeito à sua função (Basri et al., 2017).

O **método** é a descrição de qualquer procedimento específico seguido para determinar um plano otimizado de manutenção preventiva sob certos requisitos ou restrições de manutenção. O objetivo é ajudar na análise relacionada através de um procedimento estabelecido e sistemático, o que pode facilitar a obtenção de certos resultados precisos e eficientes. Assim, a tomada de decisão no planeamento de manutenção preventiva baseia-se nos métodos aplicados. Os métodos aplicados são: Inteligência artificial, Simulação, formulação matemática, formação de matrizes, Análise crítica, Multicritérios (Basri et al., 2017).

### 2.2.2 Manutenção Preditiva

A manutenção preditiva é uma metodologia de manutenção de equipamentos e máquinas em que a manutenção é programada com base em previsões de quando o equipamento é suscetível de falhar, em oposição à manutenção reativa tradicional ou à manutenção preventiva baseada no tempo. A manutenção preditiva utiliza várias técnicas, tais como a monitorização de condições, para analisar dados de sensores ou outros equipamentos, para poder prever quando é necessária a manutenção. O objetivo da manutenção preditiva é minimizar o tempo de paragem do equipamento e prolongar a sua vida útil, enquanto reduz os custos de manutenção («Contents», 2002).

Uma das principais técnicas utilizadas na manutenção preditiva é a monitorização da condição, que envolve a utilização de sensores ou outro equipamento para recolher dados sobre a condição do equipamento, tais como vibração, temperatura e análise do óleo (Cakir et al., 2021). Estes dados são então analisados utilizando vários métodos, tais como análise estatística aplicada à inteligência artificial, para identificar padrões que indicam quando é provável que o equipamento falhe.

Outra técnica utilizada na manutenção preditiva é a análise de modo de falha e efeitos, que é um processo de identificação de potenciais modos de falha do equipamento e de avaliação do risco dessas falhas. Este processo ajuda a identificar e priorizar áreas de equipamento com maior probabilidade de falha, de modo que a manutenção possa ser focada nessas áreas.

A manutenção preditiva pode ser aplicada a uma vasta gama de equipamentos, incluindo maquinaria industrial, sistemas [Aquecimento, Ventilação e Ar Condicionado \(AVAC\)](#), e equipamento de transporte. A manutenção preditiva pode ser utilizada em várias indústrias, tais como a indústria transformadora, aeroespacial, petróleo e gás, automóvel, e muitas outras (Bousdekis et al., 2020). Os benefícios da manutenção preditiva incluem:

- Reduzido tempo de paragem;
- Prolongamento da vida útil do equipamento;
- Custos de manutenção reduzidos;
- Segurança melhorada;
- Qualidade de produção melhorada;

Contudo, a implementação da manutenção preditiva pode ser um processo complexo e demorado, exigindo a utilização de software e equipamento especializado, bem como de pessoal treinado para interpretar os dados e efetuar a manutenção necessária. Além disso, os dados recolhidos devem ser analisados e interpretados corretamente, para fazer previsões exatas, o que requer formação e experiência.

### 2.2.3 Manutenção Corretiva

A manutenção corretiva é um tipo de manutenção que é efetuada em equipamentos ou sistemas depois de estes terem falhado ou já não funcionarem corretamente. O objetivo da manutenção corretiva é restaurar o equipamento ou sistema à sua condição normal de funcionamento (Y. Wang et al., 2014), e normalmente envolve a reparação ou substituição de componentes avariados.

A manutenção corretiva pode ser executada utilizando uma variedade de técnicas, incluindo a resolução de problemas, a análise das causas de raiz, e o modo de falha e análise de efeitos. A resolução de problemas é o processo de identificação da causa de uma falha e a determinação da ação corretiva adequada para a sua reparação. A análise da causa raiz é um método mais profundo que envolve a investigação das causas subjacentes a uma falha, a fim de evitar que esta se repita no futuro. A Análise de Modos de Falha e Efeitos é um processo de identificação de potenciais modos de falha do equipamento e de avaliação do risco dessas falhas (Schneider, 1996).

Novas estratégias de manutenção corretivas são importantes para assegurar um correto funcionamento, enquanto minimizam os tempos de paragem. Características tais como deteção remota de falhas, previsão de falhas e estratégias de manutenção baseadas na condição do equipamento precisam de ser incorporadas nos sistemas de monitorização para resolver inteligentemente os problemas durante o funcionamento dos sistemas (Livera et al., 2021).

Em conclusão, a manutenção corretiva é um tipo de manutenção que é realizada em equipamento ou sistemas depois de estes terem falhado, o seu objetivo é restaurar o equipamento ou sistema ao seu estado normal de funcionamento. Normalmente envolve a reparação ou substituição de componentes avariados, e pode ser executada utilizando uma variedade de técnicas como a resolução de problemas, análise da causa raiz, e modo de falha e análise de efeitos. Contudo, a manutenção corretiva é tipicamente mais dispendiosa e demorada do que a manutenção preventiva ou preditiva, uma vez que envolve frequentemente reparações e substituição de peças mais extensas, e pode levar a paragens inesperadas.

### 2.2.4 Manutenção Baseada na Condição

A **Manutenção com Base na Condição (MBC)** é um método de realizar a manutenção no equipamento que se baseia na condição real do equipamento, em vez de se basear num calendário definido ou após a ocorrência de uma falha. A monitorização baseada na condição é o processo de observar um conjunto de parâmetros e/ou variáveis que indicam o estado do sistema sob investigação. Desempenha um papel significativo na manutenção, gestão, e operações sustentáveis de vários sectores, tais como, indústrias transformadoras, transportes, energia, recursos naturais, recursos naturais, catástrofes naturais e cuidados de saúde. Na maioria dos casos, os sensores e/ou microcontroladores são amplamente utilizados para realizar a monitorização das condições (Ali & Abdelhadi, 2022).

A implementação da MBC pode ser encontrada em várias aplicações, tais como modelação de vibrações, desempenho de sensores, processamento de sinais, controlo de ruído, monitorização do desempenho termodinâmico, óleo lubrificante, monitorização da corrosão. A seleção de uma MBC apropriada (e as suas técnicas de otimização) é sempre um desafio e depende muito do sistema (Ali & Abdelhadi, 2022).

A manutenção baseada na condição é geralmente associada com manutenção preditiva, mas a sua diferença é que a manutenção preditiva utiliza dados históricos e algoritmos de ML para prever falhas no equipamento, enquanto a MBC depende da monitorização em tempo real do equipamento para detetar alterações no seu estado.

## 2.3 Manutenção Preditiva com Inteligencia Artificial

Nesta seção, foi conduzida uma análise aprofundada da manutenção preditiva, com ênfase no uso de inteligência artificial. Foi realizada uma revisão dos métodos empregados nessa abordagem, bem como foram fornecidos exemplos de aplicações práticas.

### 2.3.1 O Conceito de Inteligência Artificial e Machine Learning e a sua Relação com a Manutenção Preditiva

A Inteligência Artificial é um ramo da informática que lida com a criação de sistemas inteligentes que funcionam e reagem como os seres humanos. Envolve o desenvolvimento de algoritmos e sistemas que podem executar tarefas que normalmente exigiriam inteligência humana, tais como a resolução de problemas, a tomada de decisões e o reconhecimento de padrões (Agrawal et al., 2016).

O ML é um subconjunto de IA que permite às máquinas aprender com os dados e melhorar o seu desempenho ao longo do tempo sem serem explicitamente programadas. Implica o uso de algoritmos e modelos estatísticos que podem analisar e aprender com os dados, identificar padrões, e fazer previsões (El Naqa & Murphy, 2015).

No contexto da manutenção preditiva, IA e ML são utilizados para analisar dados de sensores e outras fontes para prever quando é provável que o equipamento falhe e programar a manutenção em conformidade. Isto pode melhorar a eficiência e eficácia das operações de manutenção, reduzindo o risco de paragens inesperadas e prolongando a vida útil do equipamento (Cardoso & Ferreira, 2021).

A utilização de IA e na manutenção preditiva envolve várias etapas, tais como recolha de dados, pré-processamento de dados, seleção de características, formação de modelos e avaliação de modelos (Cardoso & Ferreira, 2021). Os dados recolhidos do equipamento são analisados pelos modelos de IA, que podem identificar padrões e tendências que indicam quando o equipamento é suscetível de falhar. Com base nesta informação, o sistema pode programar a manutenção de forma proactiva, em vez de reativa, para que a manutenção possa ser executada antes da ocorrência de uma falha (Cardoso & Ferreira, 2021).

### 2.3.2 Benefícios do Uso de IA para a Manutenção Preditiva

Há vários benefícios em utilizar IA na manutenção preditiva que incluem (K. Wang & Wang, 2018):

- **Maior exatidão:** Os algoritmos de IA podem analisar grandes quantidades de dados e identificar padrões e tendências que são difíceis de detetar pelos humanos. Isto pode resultar em previsões mais exatas de quando é provável que o equipamento falhe, reduzindo o risco de paragens inesperadas;
- **Tomada de decisão mais rápida:** A IA podem processar dados em tempo real e fazer previsões quase instantaneamente, permitindo uma tomada de decisão mais rápida e um planeamento de manutenção mais eficaz.
- **Automatização de tarefas repetitivas:** A IA podem automatizar tarefas repetitivas e demoradas, tais como recolha de dados, análise, e relatórios, o que pode poupar tempo e recursos;
- **Prolongamento da vida útil do equipamento:** Ao prever quando é provável que o equipamento falhe e programar a manutenção em conformidade, a IA podem ajudar a prolongar a vida útil do equipamento, detetando os problemas mais cedo;
- **Redução de custos:** Ao reduzir o tempo de paragem, prolongar a vida útil do equipamento, e automatizar tarefas repetitivas, a IA e o ML podem ajudar a reduzir os custos de manutenção e aumentar a eficiência global do equipamento;

### 2.3.3 Tipos de Modelos Usados de IA para a Manutenção Preditiva

O método que tem vindo a receber maior atenção nos últimos anos é o prognóstico baseado em dados. O sistema de manutenção preditiva orientado por dados consiste em duas fases: primeiro, é necessário um processo de aprendizagem (isto é, formação de modelos) com base em sinais históricos de sensores brutos; segundo, o modelo treinado é aplicado para prever objetivos e tomar decisões. Em geral, cada fase consiste nos três subprocessos seguintes:

1. Aquisição e pré-processamento de dados, que podem ser uni-sensoriais ou multi-sensoriais;
2. Engenharia de características, que contém extração, concatenação, e seleção de características;
3. Na fase de formação do modelo para previsão, utilizam-se dados históricos para treinar um modelo, otimizando os seus hiperparâmetros. Esse modelo é adaptado para prever eventos futuros, como falhas em sistemas, com base em padrões extraídos dos dados.

Depois disso, o modelo pode prever o fluxo de dados em tempo real (W. Zhang et al., 2019). Existem várias técnicas e modelos utilizados em IA para a manutenção preditiva, que incluem:

#### Árvore de Decisão:

As árvores de decisão são um tipo de algoritmo de aprendizagem supervisionada que podem ser utilizadas tanto para tarefas de classificação como de regressão. Elas são uma representação gráfica de um modelo de decisão que divide os dados em subconjuntos com base em características específicas. Esses modelos são não paramétricos, o que significa que não fazem suposições sobre a distribuição dos dados, e são altamente interpretáveis, o que os torna uma escolha popular em diversas aplicações (Piramuthu et al., 1998).

No contexto da manutenção preditiva, as árvores de decisão podem ser utilizadas para prever o estado do equipamento (por exemplo, saudável ou defeituoso) com base nos dados históricos das leituras dos sensores. O algoritmo da árvore de decisão funciona dividindo recursivamente os dados em subconjuntos mais pequeno, com base nos valores das características de entrada como podemos ver pela Figura 2.2. Cada partição é representada por um nó na árvore, e a partição final representa uma previsão ou uma decisão (Kaparathi & Bumblauskas, 2020) (Ouadah et al., 2022).

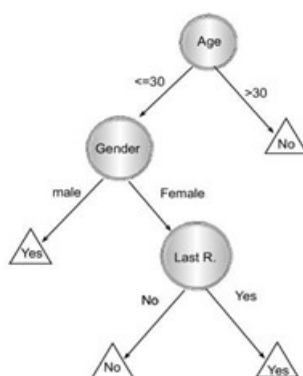


Figura 2.2: Estrutura Árvore de Decisão; Retirada de Rokach e Maimon (2005).

Começando no nó raiz, o algoritmo avalia as características disponíveis e escolhe aquela que resulta na divisão mais eficaz dos dados, geralmente usando critérios como a impureza de Gini. Em problemas de classificação binária, avalia a probabilidade de uma classificação incorrecta quando a um ponto de dados selecionado aleatoriamente é atribuída uma etiqueta de classe com base na distribuição das classes num determinado nó. Varia de 0 a 0,5, em que 0 indica um nó perfeitamente puro (todas as instâncias pertencem à mesma classe) e 0,5 significa impureza máxima (uma distribuição igual de classes). A cada nó, a árvore cria ramos que representam características ou condições, e os dados são divididos em subconjuntos cada vez mais homogêneos à medida que descem pela árvore. O processo continua até que critérios de parada sejam atingidos, como uma profundidade máxima da árvore. Para classificar ou prever novos dados, eles percorrem a árvore da raiz às folhas, seguindo os ramos correspondentes às características de entrada,

e a decisão final é obtida em um nó de folha. Podemos ver o cálculo de impureza de Gini pela Equação 2.1, para decidir qual a característica a dividir em cada nó (Kaparthi & Bumblauskas, 2020).

$$Gini(p) = 1 - (p^2 + (1 - p^2)) \quad (2.1)$$

- Gini(p) é o índice de Gini para um problema de classificação binária;
- p representa a probabilidade de um elemento pertencer a uma classe (por exemplo, "Classe A" em contraste com "Classe B").

A probabilidade "p" em relação à fórmula do índice de Gini para classificação binária pode ser obtida através dos dados de treinamento do modelo. É possível calcular a probabilidade "p" mediante a contagem do número de amostras que pertencem à classe positiva (por exemplo, "Classe A") e, em seguida, dividir esse valor pelo número total de amostras no conjunto de treinamento. A fórmula que expressa "p" é a seguinte:

$$p = \frac{\text{Número de amostras da Classe A}}{\text{Número total de amostras no conjunto de treinamento}} \quad (2.2)$$

Uma das vantagens de utilizar a árvore de decisão na manutenção preditiva é que elas podem lidar com dados contínuos e categóricos e podem modelar relações não lineares entre as características de entrada e o resultado. Além disso, são fáceis de interpretar e compreender, o que as torna úteis para explicar o raciocínio por detrás das previsões. Contudo, as árvores de decisão podem ser propensas a [overfitting](#), especialmente quando a árvore se torna demasiado profunda, e podem ser sensíveis a pequenas alterações nos dados (Ouadah et al., 2022).

#### **Floresta Aleatória:**

A floresta aleatória é um algoritmo de ML que pode ser usado tanto para tarefas de classificação como de regressão. Baseia-se na ideia de criar múltiplas árvores de decisão e depois combinar as suas previsões para melhorar a exatidão e robustez geral do modelo como pode ser observado pela Figura 2.3 (Ouadah et al., 2022).

O algoritmo funciona criando múltiplas árvores de decisão utilizando subconjuntos aleatórios dos dados e características, e depois calculando a média ou votando sobre as previsões das árvores individuais (Ouadah et al., 2022).

1. Se o número de casos no conjunto de formação for N, então a amostra de N casos é colhida aleatoriamente, mas com substituição. Esta amostra será o conjunto de treino para o cultivo da árvore;
2. Supondo que existem "M" variáveis de entrada, um número "m" é especificado de modo que em cada nó, "m" variáveis sejam selecionadas aleatoriamente de "M", e a melhor alocação sobre este "m" é utilizada para dividir o nó. O valor de "m" é mantido constante à medida que a floresta cresce;



3. Cada árvore é cultivada na maior extensão possível. Não há tamanho.

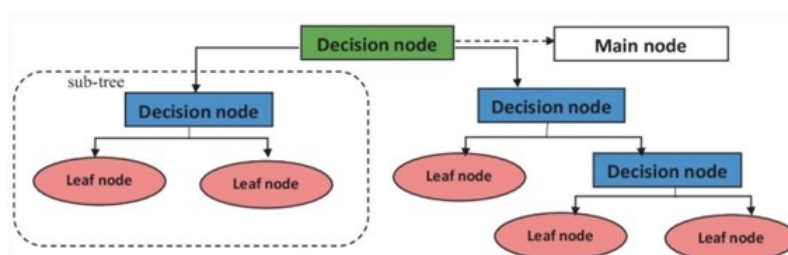


Figura 2.3: Estrutura da Floresta Aleatória; Retirada de Ouadah et al. (2022).

**Support Vector Machine (SVM):** A SVM é a mais recente técnica de aprendizagem supervisionada por máquinas. Os modelos SVM estão intimamente relacionados com as redes neurais. As SVM giram em torno da noção de "margem" em ambos os lados de um **hiperplano** que separa duas classes de dados como pode ser observado pela Figura 2.4. A maximização da margem e, assim, a criação da distância mais significativa entre o hiperplano separador e as instâncias de ambos os lados provou reduzir um limite superior sobre o erro de generalização esperado (Ouadah et al., 2022).

Isso é alcançado ao encontrar o hiperplano de separação ótimo, que fornece a maior margem possível entre as classes de dados, permitindo assim uma classificação mais precisa. Para encontrar esse hiperplano ótimo, as SVM utilizam técnicas de otimização matemática, como a maximização da margem e a minimização do erro de classificação, resultando em modelos eficazes em tarefas de classificação.

Geralmente, as SVMs e as redes neurais tendem a ter um desempenho muito melhor quando se trata de multidimensionais e características contínuas. Por outro lado, os sistemas baseados na lógica tendem a ter um melhor desempenho quando se trata de características discretas/categóricas. Para modelos de redes neurais e SVMs, é necessário um grande tamanho de amostra para atingir a sua máxima exatidão de previsão (Osisanwo et al., 2017).

É importante notar que as SVMs são sensíveis à escolha da **função Kernel** e aos valores do parâmetro de regularização, que controla o compromisso entre maximizar a margem e minimizar o erro de classificação (Ouadah et al., 2022). Além disso, pode ser computacionalmente caro para grandes conjuntos de dados, e não é adequado para conjuntos de dados com muitas características.

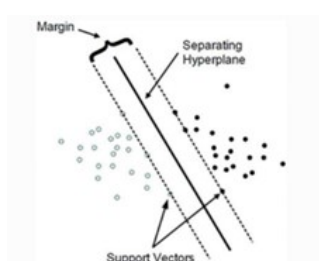


Figura 2.4: Máquina de Vetores de Suporte; Retirada de Ouadah et al. (2022).



**Artificial Neural Network (ANN):**

As redes neurais são um tipo de algoritmo de aprendizagem que se inspiram na estrutura e função do cérebro humano (Alpaydin, 2020). Consistem em camadas de nós interligados como se pode ver pela Figura 2.5, chamados neurónios, que são utilizados para modelar relações complexas entre as características de entrada e de saída (Murat & Ceylan, 2006). As redes neurais podem ser utilizadas para uma variedade de tarefas, incluindo reconhecimento de imagem, processamento de linguagem natural e previsão (Esen et al., 2008).

No contexto da manutenção preditiva, as redes neurais podem ser usadas para analisar dados de sensores e prever o estado de saúde do equipamento (Zonta et al., 2020). Podem ser treinadas com dados históricos dos sensores para aprender os padrões de comportamentos normais e anormais do equipamento. Uma vez formada, a rede neuronal pode então ser utilizada para prever o estado de saúde do equipamento em tempo real, através da análise dos dados atuais dos sensores (Zonta et al., 2020).

Existem vários tipos de redes neurais que podem ser utilizadas para manutenção preditiva, dependendo do tipo de dados e da tarefa específica em questão, conforme discutido por K. Wang e Wang (2018) existem três principais abordagens. Por exemplo:

- **Rede Neurais Feedforward (do inglês, *Feedforward Neural Network (FNN)*):** são um tipo de rede neurais que consiste numa camada de entrada, uma ou mais camadas ocultas, e uma camada de saída. São adequadas para tarefas como a classificação e a regressão (Krogh, 2008);
- **Rede Neurais Recorrentes (do inglês, *Recurrent Neural Network (RNN)*):** são um tipo de rede neuronal que é concebida para processar sequências de dados, tais como séries cronológicas ou linguagem natural. São particularmente úteis para tarefas tais como tradução de línguas, reconhecimento de fala, geração de texto, e previsão de sequências (Dixit & Silakari, 2022);
- **Redes Neurais Convolucional (do inglês, *Convolutional Neural Network (CNN)*):** são um tipo de rede neuronal que é particularmente adequada para tarefas de processamento de imagem e vídeo. São compostas por várias camadas de camadas convolutas e de agrupamento, seguidas por camadas totalmente ligadas (O'Shea & Nash, 2015);

É importante notar que as redes neurais são modelos complexos que podem exigir grandes quantidades de dados e recursos computacionais para a sua formação. Além disso, podem ser sensíveis à escolha de **hiperparâmetros**, tais como o número de camadas e o número de neurónios na rede. Além disso, é importante validar o desempenho dos modelos da rede neuronal utilizando técnicas apropriadas, como a validação cruzada, para assegurar que os modelos generalizem bem os dados "invisíveis" (K. Wang & Wang, 2018). Esses dados são chamados de "invisíveis" porque o modelo não teve acesso a eles durante o processo de treinamento.

Além disso, é importante pré-processar os dados de forma apropriada para assegurar que a rede neuronal possa extrair as características relevantes dos dados. É também importante prestar atenção à interpretabilidade e explicabilidade das redes neurais, compreender como o modelo está a fazer as suas previsões e identificar os fatores mais importantes na determinação da saúde do equipamento.

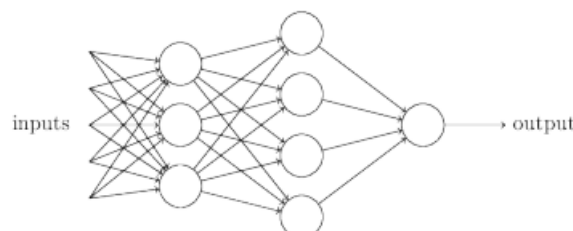


Figura 2.5: Exemplo de Rede Neuronal Artificial; Retirada de Nielsen (2015)

Como foi dito anteriormente as RNNs são um tipo de rede neuronal que são concebidas para processar dados sequenciais, tais como séries cronológicas ou texto (Chung et al., 2015). As RNNs operam em sequências de dados de entrada, mantendo um estado interno que depende das entradas anteriores. O estado interno do RNN pode ser pensado como um resumo das entradas anteriores, e pode ser utilizado para informar o processamento de entradas futuras (Schmidt, 2019).

A formação dos RNNs é deduzida a partir do gradiente descendente e da regra da cadeia. Durante a retro propagação, o gradiente da função de perda relativa à saída da rede é calculado, e depois este gradiente é retro propagado através da rede para atualizar os pesos. A regra da cadeia é utilizada para calcular o gradiente de perda para o estado interno da rede, e este gradiente é utilizado para atualizar o estado interno da rede durante a formação (A. Zhang et al., 2021).

Um dos principais problemas com a formação dos RNNs é o problema dos gradientes que desaparecem ou explodem (Chung et al., 2015). Isto ocorre quando os gradientes se tornam muito pequenos ou muito grandes à medida que se propagam através da rede durante a retro propagação. Quando os gradientes se tornam demasiado pequenos, a rede torna-se difícil de treinar, já que as atualizações dos pesos se tornam insignificantes. Pelo contrário, quando os gradientes se tornam demasiado grandes, a rede pode tornar-se instável, uma vez que as atualizações dos pesos podem causar grandes oscilações na função de perda (A. Zhang et al., 2021).

No contexto das redes neurais, o gradiente é utilizado para atualizar os pesos da rede durante o processo de formação. Os pesos são ajustados na direção do gradiente negativo da função de perda, o que diz à rede como alterar a sua saída para melhor corresponder à saída alvo.

As redes *Long Short Term Memory* (LSTM) são um tipo de RNN que se destina a aliviar o problema dos gradientes que desaparecem ou explodem. As LSTMs utilizam um mecanismo de portões que permite à rede atualizar seletivamente ou esquecer o estado interno, com base nos dados de entrada. Este mecanismo pode ajudar a evitar que os

gradientes se tornem demasiado grandes ou demasiado pequenos, e pode melhorar a capacidade da rede de capturar dependências a longo prazo nos dados de entrada (A. Zhang et al., 2021).

**Long Short-Term Memory (LSTM):** LSTM é um tipo de RNNs concebida para processar sequências de dados, tais como séries cronológicas ou linguagem natural (Rivas et al., 2020). São particularmente úteis para tarefas tais como tradução de línguas, reconhecimento de fala, geração de texto, e previsão de sequências (Smagulova & James, 2019). No contexto da manutenção preditiva, as LSTMs podem ser utilizadas para analisar os dados de séries temporais obtidos a partir de leituras de sensores para prever o estado de saúde do equipamento.

As LSTMs têm uma célula de memória que pode armazenar informação durante longos períodos, permitindo-lhes manter o contexto e recordar informação passada. Têm também portas que controlam o fluxo de informação dentro e fora da célula de memória, o que lhes permite esquecer ou recordar seletivamente a informação (Smagulova & James, 2019). Isto torna as LSTMs adequadas para tarefas que requerem memória a longo prazo e a capacidade de recordar informação anterior, tais como sequências de leituras de sensores ao longo do tempo.

O modelo LSTM é um tipo de rede neuronal recorrente que utiliza um conjunto de três portões para aprender e recordar seletivamente a informação da sequência de entrada. A porta de entrada determina quanta informação nova a acrescentar ao estado da célula, a porta de esquecimento determina quanto do estado da célula anterior a reter, a porta de saída determina quanto do estado atualizado da célula a sair, e o estado da célula representa a memória da célula LSTM. Ajustando cuidadosamente os pesos e enviesamentos destas portas durante o treino, o modelo LSTM pode aprender a processar e prever sequências de dados com alta exatidão.

Em manutenção preditiva, uma rede LSTM pode ser treinada com dados históricos de sensores para aprender os padrões de comportamentos normais e anormais do equipamento. Uma vez treinada, a rede pode então ser utilizada para prever o estado de saúde do equipamento em tempo real, através da análise dos dados atuais dos sensores (Smagulova & James, 2019).

As LSTMs podem ser treinadas de diferentes formas, dependendo dos tipos de dados disponíveis. Por exemplo, as LSTMs univariável são usadas para dados de séries temporais de um único sensor, enquanto as LSTMs multivariável lidam com dados de vários sensores. Além disso, a aprendizagem por transferência permite otimizar LSTMs pré-treinadas para tarefas e conjuntos de dados específicos. Esta técnica envolve aplicar o conhecimento de uma tarefa para melhorar o desempenho noutra tarefa relacionada, mesmo que em domínios diferentes. A justificação para isso é que os modelos pré-treinados já possuem conhecimento prévio e a capacidade de identificar padrões complexos nos dados. A adaptação desses modelos a tarefas e dados específicos aproveita esse conhecimento prévio e a capacidade de generalização, resultando em melhores resultados.

É importante notar que a LSTM é uma técnica complexa, e pode requerer grandes

quantidades de dados e recursos computacionais para treinar. O modelo pode ser sensível à escolha de hiperparâmetros, tais como o número de camadas e o número de neurónios na rede. Além disso, é importante validar o desempenho do modelo LSTM utilizando técnicas apropriadas, como a validação cruzada, para assegurar que o modelo generaliza bem os dados “invisíveis”.

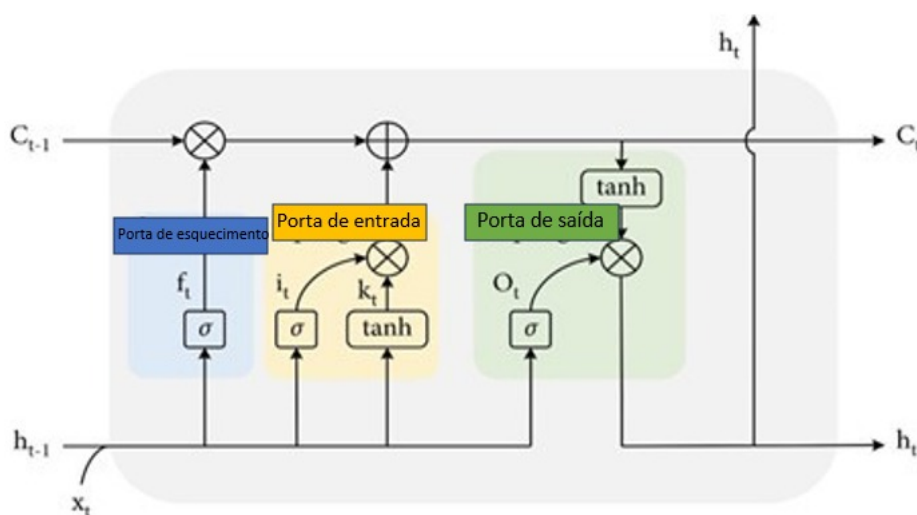


Figura 2.6: Esquema de Funcionamento e Classificação dos Gates; Retirada de Zbiciak e Markiewicz (2023)

O modelo LSTM consiste numa série de células LSTM, cada uma das quais contém três componentes principais ou "portas": a porta de entrada, a porta de esquecimento, a porta de saída como podemos ver pela Figura 2.6. Cada portão é controlado por um conjunto de pesos e *bias* que são aprendidos durante o treino usando o gradiente descendente e regra da cadeia. Há três componentes adicionais cruciais que podem ser observados nas linhas representadas na Figura 2.7, A linha verde na figura representa a memória a longo prazo, também conhecida como o estado da célula. Ao entrar na célula LSTM, é referida como o estado de célula anterior. Após passar por modificações, será denominada simplesmente estado de célula.

Além disso, há a linha cor-de-rosa que corresponde à memória de curto prazo, também chamada de estado oculto. Semelhante à linha verde, quando entra na célula LSTM, é designada como estado oculto anterior. Depois de sofrer modificações, será simplesmente chamada estado oculto. Temos ainda a linha cinzenta que representa o input (A. Zhang & Lipton, 2023).

A porta do esquecimento é responsável por determinar a quantidade de memória de longo prazo que deve ser mantida. Para fazer isso, ela recebe como entrada o estado oculto anterior, com seus pesos, e o *input*, também com seus pesos. Depois, soma esses dois valores e adiciona o *bias* (Equação 2.3). O resultado é então passado pela função sigmoide, que produz um valor entre 0 e 1 (Equação 2.4). Esse valor indica a percentagem do estado da célula anterior que deve ser mantida (A. Zhang & Lipton, 2023).

A porta de entrada, por sua vez, determina a quantidade de memória potencial de longo prazo que deve ser adicionada ao estado da célula anterior. Ela é composta por dois blocos. O bloco da esquerda recebe o estado oculto anterior e o *input*, soma os dois valores e adiciona o *bias* (Equação 2.5). O resultado é passado pela função sigmoide, que produz um valor entre 0 e 1 (Equação 2.6). Esse valor indica a percentagem que será utilizada do bloco da direita. O bloco da direita recebe o estado oculto anterior e o *input* com os seus respectivos pesos, soma os dois valores e adiciona o *bias* (Equação 2.7). O resultado é passado pela função tanh, que produz um valor entre -1 e 1 (Equação 2.8). Esse valor é a memória potencial de longo prazo. Os dois blocos são multiplicados, produzindo a percentagem de memória potencial de longo prazo (Equação 2.9) (A. Zhang & Lipton, 2023).

Por fim, a porta de saída é responsável por atualizar o estado oculto anterior. Ela também é composta por dois blocos. O bloco da direita recebe o estado da célula atualizado e o *input*, soma os dois valores e adiciona o *bias* (Equação 2.12). O resultado é passado pela função sigmoide, produzindo um valor entre 0 e 1 (Equação 2.13), que indica a percentagem que será utilizada do bloco da esquerda. Os dois blocos são multiplicados, produzindo o estado atualizado do estado oculto da célula e o *output* do valor da célula LSTM Equação 13 (A. Zhang & Lipton, 2023).

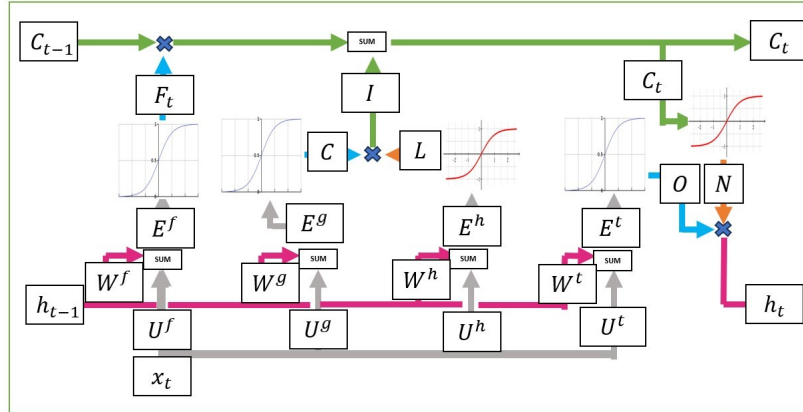


Figura 2.7: Ilustração dos Processos Matemáticos de uma Célula LSTM.

$$X_f = (h_{(t-1)} \times W^f + x_t \times U^f) + E^f \quad (2.3)$$

$$F_{(t)}(X_f) = \frac{e^x}{e^x + 1} \quad (2.4)$$

$$X_g = (h_{(t-1)} \times W^g + x_t \times U^g) + E^g \quad (2.5)$$

$$\dot{C}(X_g) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.6)$$

$$X_h = (h_{(t-1)} \times W^h + x_t \times U^h) + E^h \quad (2.7)$$

$$L(X_h) = \frac{e^x}{e^x + 1} \quad (2.8)$$

$$I = L(X_g) \times \dot{C}(X_h) \quad (2.9)$$

$$C_{(t)} = C_{(t-1)} \times F_{(t)} + I \quad (2.10)$$

$$N(C_{(t)}) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.11)$$

$$X_b = (h_{(t-1)} \times W^t + x_t \times U^t) + E^t \quad (2.12)$$

$$O(X_b) = \frac{e^x}{e^x + 1} \quad (2.13)$$

$$h_{(t)} = O \times N \quad (2.14)$$

$h_{(t-1)}$ : Estado Oculto Anterior;

$h_{(t)}$ : Estado Oculto Atual;

$C_{(t-1)}$ : Estado anterior da célula;

$C_{(t)}$ : Estado da célula Atual;

$F_{(t)}$ : Percentagem de Estado anterior da célula deve ser retido para o passo de tempo atual;

$\dot{C}$ : Percentagem de importância para o passo de tempo atual;

$L$ : Memória potencial de longo prazo;

$I$ : Percentagem de vetor candidato que é retido para o estado anterior da célula;

$O$ : Percentagem do estado oculto anterior que deve ser produzido na etapa de tempo atual;

$N$ : Memória potencial de curto prazo

$x_t$ : Vetor de Entrada Atual;

$W^f, W^g, W^h, W^t$ : Pesos de Estado Oculto Anterior em cada entrada;

$U^f, U^g, U^h, U^t$ : Pesos de Vetor de Entrada Atual em cada entrada;

$E^f, E^g, E^h, E^t$ : Bias de Vetor de Entrada Atual em cada entrada;

$X_b, X_h, X_f$ : Valores de Entrada nas Funções sigmoide;

$X_g$ : Valores de Entrada na Função tanh.

**Convolutional Neuronal Networks (CNNs):**

CNNs são um tipo de algoritmo de aprendizagem profunda que é particularmente adequado para tarefas de processamento de imagem e vídeo (Bartoli & Fusiello, 2021). São inspiradas pelo sistema de processamento visual do cérebro humano e são compostas

por múltiplas camadas de camadas convolutas e de agrupamento, seguidas por camadas totalmente ligadas como podemos ver pela Figura 2.8.

No contexto da manutenção preditiva, as CNNs podem ser utilizadas para analisar imagens obtidas de sensores como câmaras ou imagens térmicas, para prever o estado de saúde do equipamento. Por exemplo, as CNN podem ser treinadas para classificar as imagens do equipamento em categorias saudáveis ou defeituosas, com base em padrões visuais, tais como fissuras ou deformações no equipamento (Xie et al., 2020).

Durante o treinamento, as CNNs aprendem a ajustar filtros para identificar características como bordas, texturas e padrões na imagem. Cada filtro é projetado para detetar uma característica específica por meio de uma operação de convolução na imagem de entrada, resultando em mapas de características que destacam as características identificadas. Posteriormente, camadas de agrupamento reduzem a dimensionalidade dos mapas de características para simplificar a representação da imagem, mantendo as informações essenciais para tarefas de classificação e detecção (Bartoli & Fusiello, 2021).

É importante notar que as CNN exigem grandes quantidades de dados rotulados para treinar e podem ser sensíveis à escolha de hiperparâmetros, tais como o número de camadas e o número de filtros na rede. Além disso, é importante validar o desempenho do modelo CNN utilizando técnicas apropriadas tais como a validação cruzada, para assegurar que o modelo generaliza bem os dados “invisíveis”. Além disso, é importante pré-processar as imagens de forma adequada para assegurar que a CNN possa extrair as características relevantes das imagens.

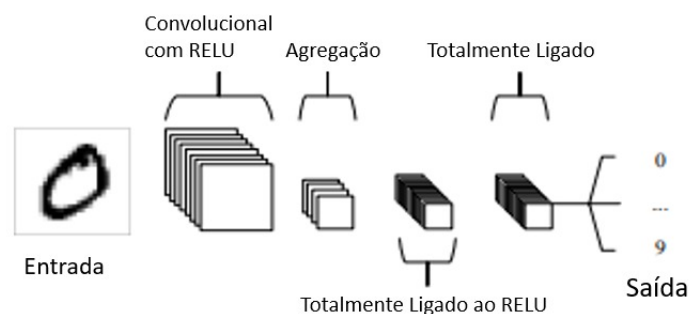


Figura 2.8: Uma arquitetura CNN simples, composta por cinco camadas; Retirada de O'Shea e Nash (2015)

### 2.3.4 Limitações à Implementação da IA na Manutenção Preditiva

Existem vários desafios e limitações à implementação da Inteligência Artificial na manutenção preditiva:

- **Falta de dados:** Um dos principais desafios na implementação da IA e do ML na manutenção preditiva é a falta de dados. A manutenção preditiva depende de uma



grande quantidade de dados históricos para formar os modelos, e se os dados não estiverem disponíveis ou não forem de boa qualidade, os modelos podem não ter um bom desempenho. Isto pode ser especialmente desafiante para equipamento novo;

- **Dificuldade em prever falhas no equipamento:** Outro desafio na implementação de IA e ML na manutenção preditiva é a dificuldade em prever falhas no equipamento. A falha do equipamento é um processo complexo que pode ser influenciado por muitos fatores, e pode ser difícil modelar as relações entre as características de entrada e o resultado. Além disso, alguns equipamentos podem ter modos de falha que não são facilmente detetáveis pelos sensores, o que pode dificultar a previsão da falha;
- **Variabilidade dos dados e ruído:** Os modelos de manutenção preditiva são treinados para reconhecer padrões nos dados dos sensores, mas estes padrões podem variar dependendo das condições do equipamento, do ambiente e dos próprios sensores. Isto pode tornar difícil generalizar os modelos a novas situações e detetar alterações subtis no comportamento do equipamento que possam indicar uma falha iminente;
- **Interoperabilidade e compreensão dos modelos:** Algumas das técnicas mais avançadas de IA e ML, tais como o *deep learning*, podem ser difíceis de interpretar e explicar. Isto torna difícil compreender como o modelo está a fazer as suas previsões e identificar os fatores que são mais importantes na determinação da saúde do equipamento;
- **Resistência à mudança:** Outro desafio da implementação de IA e ML na manutenção preditiva é a resistência à mudança. Pode haver resistência à implementação de novas tecnologias, especialmente se as práticas de manutenção existentes estiverem a funcionar bem, ou se a pessoa responsável pela manutenção não estiver familiarizada com as novas tecnologias;

É importante notar que estes desafios podem ser mitigados utilizando técnicas apropriadas tais como o pré-processamento de dados, seleção de características, e validação de modelos, e envolvendo o pessoal responsável pela manutenção no processo de implementação da IA e do ML (W. Zhang et al., 2019).

### 2.3.5 Aplicação da Inteligência Artificial na Manutenção Preditiva: Um Estudo de Caso

Este estudo utilizou um processo abrangente em três fases para o desenvolvimento de um algoritmo robusto de deteção de falhas. O processo teve início com a aquisição e pré-processamento de dados, seguido pelo processamento do modelo e culminando com o pós-processamento.



Para dar início ao estudo, a fase de pré-processamento de dados mostrou-se de extrema importância, sobretudo devido às complexas condições operacionais das turbinas eólicas. Estas condições estavam caracterizadas por restrições de energia e pela presença de valores atípicos nos dados históricos do SCADA. A resolução destes desafios revelou-se de importância crucial, e o estudo adotou metodicamente uma abordagem sistemática.

Durante a fase de pré-processamento de dados, o estudo empregou o método da floresta de isolamento, um algoritmo de aprendizagem automática apropriado para a detecção de valores atípicos. Este método da floresta de isolamento foi concebido com o propósito de identificar instâncias de anomalia e de identificar os atributos críticos responsáveis pela degradação do equipamento. Para atingir tal objetivo, os dados foram aleatoriamente particionados e árvores de isolamento foram criadas. O algoritmo, subsequentemente, quantificou o comprimento médio do caminho para cada ponto de dados, o que, por sua vez, determinou a sua pontuação de anomalia. Os pontos de dados com comprimentos médios de trajeto mais curtos foram considerados como anómalos, indicando desvios em relação ao comportamento esperado ou normal.

Após a fase de pré-processamento de dados, enriquecida pelo conhecimento do método da floresta de isolamento, os dados limpos foram posteriormente introduzidos em três algoritmos de aprendizagem automática distintos. Estes algoritmos compreendiam a Regressão Linear Múltipla, o Reforço de Gradiente Extremo (XGBoost) e o LSTM. A fase crucial do processamento do modelo envolveu um processo rigoroso de seleção do modelo ideal. Esta seleção baseou-se em uma avaliação minuciosa de métricas de desempenho, tais como o R-quadrado, o erro quadrático médio (RMSE), o erro absoluto médio (MAE) e o erro percentual absoluto médio (MAPE).

Um destaque particular recai sobre a consideração das variáveis de entrada para a previsão de falhas. Para a previsão da "Temperatura do rolamento da caixa de velocidades", foram levadas em conta variáveis-chave, como a temperatura da nacela, a velocidade do rotor, a potência ativa, a temperatura externa e a temperatura do cárter de óleo da caixa de velocidades. Estas variáveis desempenharam um papel fundamental na previsão do estado da caixa de velocidades.

De forma semelhante, para a previsão da "Temperatura do rolamento do gerador", o modelo considerou a temperatura da nacela, a potência ativa, a velocidade do gerador e a temperatura do estator do gerador. Estas entradas foram criteriosamente selecionadas para retratar de forma precisa o estado do sistema do gerador.

O conjunto de dados foi dividido em dois segmentos: um conjunto de treino e um conjunto de teste. Esta divisão implicou a alocação das 70% iniciais dos dados para fins de treino e a reserva dos 30% finais para rigorosos testes.

O ponto culminante do estudo envolveu o pós-processamento da variável de saída prevista. Este processo incluiu a avaliação de desvios em relação aos registos históricos reais. Adicionalmente, foram estabelecidos limiares críticos de falha para cada modelo em consideração. Os pontos de dados que ultrapassaram esses limiares, juntamente com as alterações nas médias, serviram como indicadores claros de potenciais falhas nas turbinas

eólicas.

### 2.4 Manutenção Preditiva em Instalações Fotovoltaicas

Neste capítulo, será abordado o tema dos benefícios da manutenção preditiva em instalações fotovoltaicas. Serão discutidos os diferentes tipos de dados que são importantes para a realização da manutenção preditiva e serão apresentados exemplos de problemas que podem ser previstos com o auxílio da inteligência artificial em parques fotovoltaicos.

A manutenção preditiva em parques fotovoltaicos é uma abordagem de manutenção proactiva que utiliza análise de dados, IA, e outras técnicas avançadas para prever quando é necessária uma manutenção antes que ocorra uma falha. Isto permite que a manutenção seja programada no momento mais adequado, evitando tempos de paragem dispendiosos e reduzindo os custos de manutenção.

A manutenção preditiva é um aspeto essencial para o funcionamento e manutenção de centrais fotovoltaicas. Algumas das razões pelas quais a manutenção preditiva é importante para as centrais fotovoltaicas. Algumas das razões pelas quais a manutenção preditiva é importante para as centrais fotovoltaicas, encontram-se destacadas nos seguintes pontos (Çinar et al., 2020):

- **Melhoria da fiabilidade do equipamento:** A manutenção preditiva permite uma monitorização proactiva do estado do equipamento fotovoltaico, que pode detetar potenciais problemas antes que estes resultem numa falha. Isto pode aumentar a fiabilidade do equipamento e reduzir o tempo de paragem;
- **Redução de custos:** A manutenção preditiva pode ajudar a identificar problemas antecipadamente, reduzindo a necessidade de reparações mais extensivas e prolongando a vida útil do equipamento. Isto pode resultar em poupanças de custos significativas em comparação com a manutenção reativa;
- **Aumento da produção de energia:** A manutenção preditiva pode identificar e resolver questões que possam estar a ter impacto no desempenho e eficiência dos módulos fotovoltaicos, o que pode resultar numa maior produção de energia e num maior retorno do investimento;
- **Segurança melhorada:** A manutenção preditiva pode ajudar a prevenir acidentes, identificando potenciais perigos e assegurando que o equipamento está a funcionar em segurança;
- **Benefícios ambientais:** A manutenção preditiva pode reduzir o número de falhas do equipamento e a quantidade de resíduos gerados, o que pode ter impactos ambientais positivos;

As técnicas mais utilizadas para o sucesso da manutenção preditiva em centrais fotovoltaicas é a recolha e análise de dados de várias fontes.

Estas fontes incluem, os inversores que são peças fundamentais nesses sistemas, responsáveis por converter a eletricidade gerada pelos painéis solares em eletricidade AC para ser injetada na rede. Dados como tensão e corrente DC/AC, potência de saída, eficiência, temperatura e frequência de saída são normalmente retirados dos inversores (De Benedetti et al., 2018).

Além disso, sensores são usados nos módulos fotovoltaicos para obter informações sobre a curva I-V (Z. Chen et al., 2017). Existem diversas abordagens para avaliar o desempenho da instalação, incluindo inspeção visual, análise de resistência de isolamento, termografia infravermelha e cálculos comparativos com a geração real. Outro fator importante que influencia o desempenho das tecnologias FV é a temperatura do módulo do painel, que diminui com altas temperaturas (Bosman et al., 2020). Isso ocorre porque os materiais semicondutores usados nos painéis FV, como o silício, têm seu comportamento elétrico afetado pelo calor. À medida que a temperatura aumenta, os elétrons, nesses materiais ganham energia térmica, tornando-se mais agitados e menos eficazes na conversão da luz solar em eletricidade.

Fatores ambientais como temperatura, radiação solar, humidade, precipitação e velocidade do vento (Bosman & Darling, 2018) também podem afetar o desempenho dos painéis solares e outras partes do sistema. Sensores como piranômetros, de temperatura, anemômetros, pluviômetros e higrógrafos podem ser utilizados para recolher dados sobre esses fatores (Silvestre et al., 2014).

Dados históricos de outras instalações semelhantes podem ser utilizados para identificar padrões e prever potenciais problemas (De Benedetti et al., 2018). Registos de atividades de manutenção passadas também podem fornecer insights sobre possíveis falhas no futuro.

Uma vez recolhidos os dados, estes podem ser analisados utilizando *machine learning* e *deep learning* e outras técnicas avançadas para identificar padrões e prever quando é necessária a manutenção. A manutenção preditiva pode ser utilizada para detetar e abordar uma vasta gama de assuntos, incluindo:

A falha do inversor pode causar paralisação do sistema e perda de rendimentos. No entanto, a manutenção preditiva pode ajudar a identificar problemas potenciais do inversor antes que resultem em falhas, permitindo que a manutenção seja agendada no momento mais adequado. Por exemplo, mudanças nos níveis de tensão ou corrente podem indicar problemas com os componentes internos do inversor. Flutuações na potência de saída podem ser um sinal de mau funcionamento do inversor (De Benedetti et al., 2018).

Os cabos e conectores de um parque fotovoltaico podem degradar-se com o tempo devido à exposição às condições atmosféricas. Isso pode levar à perda de isolamento e exposição, resultando em perdas de energia ou até mesmo em incêndios, curtos-circuitos ou circuitos abertos. No entanto, a manutenção preditiva pode identificar possíveis falhas nos cabos e conectores, permitindo uma manutenção ou substituição proativa. A análise dos dados de tensão e corrente pode ajudar a detetar esses problemas por meio de níveis elevados de corrente (Z. Chen et al., 2017).

Os módulos fotovoltaicos podem apresentar falhas ao longo do tempo, como hotspots, fissuras e delaminação. Isso pode reduzir a saída de energia e até mesmo levar à falha do módulo. No entanto, a manutenção preditiva pode detetar sinais precoces de falha do módulo e permitir a manutenção preventiva ou a substituição. A manutenção preditiva pode utilizar técnicas como a termografia infravermelha e leitura da curva I-V para detetar hotspots, que podem indicar defeitos nas células do módulo. Inspeções visuais com *drones* e medições elétricas também podem ser utilizadas para identificar fissuras, delaminação e outros defeitos (Z. Chen et al., 2017).

Os painéis fotovoltaicos são tipicamente organizados em strings e em paralelo, com vários módulos ligados em série. Se um ou mais módulos dentro de uma string falharem, isso pode reduzir o rendimento geral do sistema. No entanto, a manutenção preditiva pode detetar falhas de *string* monitorizando mudanças na tensão e corrente de (Silvestre et al., 2014).

#### Exemplos de Casos de Estudos de Manutenção Preditiva em Parques Fotovoltaicos de Falhas

Foi apresentada uma metodologia para prever falhas em inversores de centrais fotovoltaicas usando dados *Supervisory Control and Data Acquisition* (SCADA) (Betti et al., 2019). Recebe como dados de entrada: a Corrente DC, Tensão DC, Potência DC, Corrente AC, Tensão AC, Potência AC, Temperatura Interna, Temperatura do Pannel, Temperatura Ambiente, Irradiância Global Horizontal e Irradiância Inclinada Global. O *Supervision Diagnosis Model* (SDM) prevê falhas genéricas com 7 dias de antecedência, com uma sensibilidade de até 95% e especificidade de quase 80%. O *Short Term Fault Prediction Model* (STFPM) também apresenta boa capacidade de previsão para classes de falhas de alta frequência até 7 dias antes dos danos. Combinando esses dois modelos como pode ser observado pela Figura 2.9, os operadores podem mudar de uma manutenção reativa para uma estratégia pró-ativa, melhorando a tomada de decisão e evitando falhas. O modelo foi testado na Roménia e Grécia e pode ser facilmente implantado em novas instalações fotovoltaicas. Os próximos passos incluem a introdução de um sistema automatizado de deteção de falhas usando imagens térmicas de drones e a integração do modelo preditivo em software de monitorização solar inteligente com gestão de intervenção e ferramentas de relatório baseadas em *business intelligence*.

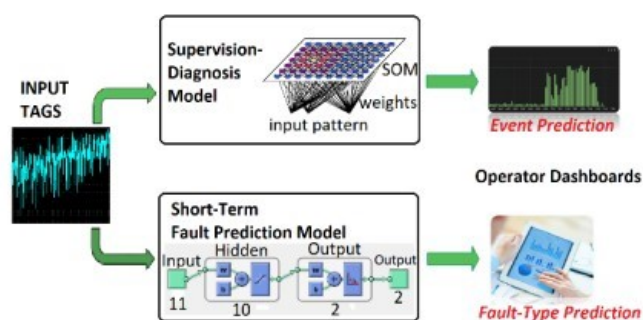


Figura 2.9: Tipos de Modelos Utilizados; Retirada de Betti et al. (2019)

Outra abordagem proposta tem como objetivo prever falhas no sistema fotovoltaico detetando anomalias antes que causem danos significativos (Benedetti, 2018). É considerado que a instalação possui um sistema adequado de recolha de dados capaz de armazenar informações relevantes sobre irradiação solar (medido por um piranómetro), temperatura e energia AC. Para a amostragem de dados, é necessário um intervalo de amostragem suficiente para produzir dados precisos com uma média de 1 hora.

A abordagem é composta por duas fases distintas. Na primeira fase, o modelo do parque/inversor que será monitorizando deve ser desenvolvido com base em dados históricos armazenados. Esse modelo será treinado usando uma rede neuronal artificial. Na segunda fase, os dados amostrados são analisados para detetar anomalias. Uma rede neuronal artificial é usada para prever a potência de saída do inversor em condições normais, e essa saída é usada para calcular a diferença diária entre as previsões e os valores reais. Valores que excedem os limites normais de funcionamento resultarão em avisos. O documento também relata resultados de dados experimentais que revelam que o modelo leva a uma boa taxa de deteção de anomalias no inversor, que é medida como uma taxa de deteção preditiva positiva superior a 90% (De Benedetti et al., 2018).

## 2.5 Inteligencia Artificial em Baterias

Nos últimos anos, a inteligência artificial tem sido aplicada às baterias com o objetivo de a tornar mais economicamente viável, devido às suas vantagens, tais como maior densidade de energia, maior vida útil, menor taxa de auto descarga e menor poluição atmosférica. A bateria de iões de lítio é o tipo mais utilizado em vários dispositivos, incluindo veículos elétricos, computadores e telemóveis. No entanto, o mau funcionamento da bateria pode levar a problemas de segurança e confiabilidade. Para monitorizar o estado da bateria em tempo real, foi conduzida uma investigação com o propósito de determinar a esperança de vida útil da bateria de lítio por meio de técnicas de IA (J. C. Chen et al., 2021). O desafio é utilizar com sucesso esses dados para prever o RUL das baterias de iões de lítio. Isso é importante para a manutenção preditiva, pois fornece indicações de aviso prévio de potencial deterioração, reduz custos de manutenção significativos e reduz a possibilidade de falhas súbitas dos dispositivos.

Este estudo (J. C. Chen et al., 2021) propõe um modelo híbrido de ciência de dados para a previsão *Remaining Useful Life* (RUL) de baterias de iões de lítio. O modelo utiliza diversas RNN, incluindo RNN clássicas, unidades recorrentes fechadas, LSTM e LSTM bidirecionais, para realizar previsões do SOH e a RUL de baterias de iões de lítio. A *otimização Bayesiana* é utilizada para encontrar os melhores hiperparâmetros de RNNs profundos. A vida útil restante total de B0005, B0006 e B0018 é de 124, 108 e 96, respetivamente. Na presente investigação (J. C. Chen et al., 2021), considera-se que a redução da capacidade da bateria de lítio para 70% constitui uma falha da bateria. Com o intuito de avaliar interactivamente o efeito do modelo, adotaram-se três categorias distintas, denominadas de *Starting Positions*

(SP)s de previsão. Para tal, definimos os SPs de 70% a 90% de SOC. Cada categoria possui momentos diferentes, com a finalidade de verificar a sua influência.

A primeira combinação consiste em utilizar 40% dos dados para treino, 30% para validação e 30% para teste. Os primeiros 70% dos dados são empregados na construção do modelo, enquanto os 30% restantes são utilizados para avaliar a eficácia do mesmo. A segunda combinação adota uma distribuição de 60% para treino, 20% para validação e 20% para teste. Neste caso, os primeiros 80% dos dados são empregados na construção do modelo, sendo os 20% restantes utilizados para testar a sua eficácia. Por fim, a terceira combinação destina 80% dos dados para treino, 10% para validação e 10% para teste. Aqui, os primeiros 80% dos dados são utilizados na construção do modelo, enquanto os 10% restantes são destinados a avaliar a eficácia obtida.

Ao ser adotada diferentes configurações, busca-se analisar o impacto dos SPs de previsão de maneira interativa, permitindo uma análise comparativa das diferentes combinações e a obtenção de resultados mais abrangentes e confiáveis para a investigação em questão. Os resultados experimentais com os dados das baterias de íões de lítio do Repositório de Dados Prognósticos Ames da NASA mostram que o modelo de ciência de dados híbrido proposto pode prever com exatidão a SOH e a RUL das baterias de íões de lítio. A rede LSTM apresenta resultados altamente satisfatórios, conforme evidenciado na Tabela 2.10, onde se destaca como o modelo mais frequente em negrito, indicando que obteve o desempenho mais consistente entre todos os modelos avaliados. Além disso, a Figura 2.11 também comprova esses resultados, demonstrando que os erros na previsão do número de ciclos de vida permanecem consistentemente abaixo de 5%. Portanto, podemos afirmar que a rede LSTM apresenta resultados notáveis. Podemos ver as métricas utilizadas na tabela pelas Equações 2.15, 2.16 e 2.17.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N}} \quad (2.15)$$

$$MAE = \frac{1}{N} \times |Predicted_i - Actual_i| \quad (2.16)$$

$$MAPE = \frac{1}{N} \times \left| \frac{Actual_i - Predicted_i}{Actual_i} \right| \quad (2.17)$$

Este estudo demonstrou a aplicação bem-sucedida da inteligência artificial e da ciência de dados na previsão do SOH e do RUL das baterias de íões de lítio. O modelo híbrido proposto, que utiliza várias RNNs profundas, revelou resultados promissores na previsão precisa do SOH e do RUL. As experiências com dados reais de baterias de íões de lítio confirmaram o desempenho satisfatório da rede LSTM. Estes resultados são relevantes para a manutenção preditiva, permitindo a detecção antecipada de falhas e a otimização dos processos de manutenção

| Battery Prediction SP Method |           |       | Evaluation Metrics |         |        |
|------------------------------|-----------|-------|--------------------|---------|--------|
|                              |           |       | RMSE               | MAE     | MAPE   |
| B0005                        | 87 (70%)  | BLSTM | 5.6562             | 3.7235  | 5.97%  |
|                              |           | LSTM  | 8.0958             | 6.9856  | 9.57%  |
|                              |           | GRU   | 8.7006             | 7.7104  | 11.17% |
|                              |           | RNN   | 5.7211             | 4.4898  | 7.11%  |
|                              | 99 (80%)  | BLSTM | 5.5567             | 4.2783  | 6.72%  |
|                              |           | LSTM  | 7.8549             | 6.9251  | 9.01%  |
|                              |           | GRU   | 6.5122             | 5.5149  | 8.34%  |
|                              |           | RNN   | 4.3688             | 3.4405  | 5.72%  |
|                              | 111 (90%) | BLSTM | 4.9642             | 4.0015  | 6.51%  |
|                              |           | LSTM  | 4.6618             | 3.8382  | 5.89%  |
|                              |           | GRU   | 5.1626             | 4.0683  | 6.83%  |
|                              |           | RNN   | 4.5135             | 3.4856  | 5.12%  |
| B0006                        | 75 (70%)  | BLSTM | 11.3130            | 9.1519  | 10.89% |
|                              |           | LSTM  | 12.4604            | 10.2055 | 12.52% |
|                              |           | GRU   | 9.9722             | 8.0417  | 8.23%  |
|                              |           | RNN   | 11.2093            | 9.3223  | 10.49% |
|                              | 87 (80%)  | BLSTM | 11.4251            | 10.8174 | 11.02% |
|                              |           | LSTM  | 12.1087            | 11.5224 | 12.01% |
|                              |           | GRU   | 10.4251            | 9.7685  | 8.77%  |
|                              |           | RNN   | 10.4307            | 9.7923  | 8.84%  |
|                              | 97 (90%)  | BLSTM | 7.7780             | 7.4168  | 7.78%  |
|                              |           | LSTM  | 5.5160             | 5.1667  | 4.89%  |
|                              |           | GRU   | 7.3028             | 7.0744  | 6.92%  |
|                              |           | RNN   | 7.7416             | 7.2983  | 7.27%  |
| B0018                        | 67 (70%)  | BLSTM | 4.8068             | 4.3762  | 5.19%  |
|                              |           | LSTM  | 3.8689             | 3.1575  | 3.87%  |
|                              |           | GRU   | 5.1798             | 4.5980  | 6.53%  |
|                              |           | RNN   | 14.6546            | 11.6006 | 12.33% |
|                              | 77 (80%)  | BLSTM | 3.5695             | 3.0946  | 4.87%  |
|                              |           | LSTM  | 1.7887             | 1.6098  | 2.26%  |
|                              |           | GRU   | 3.4913             | 2.7309  | 3.51%  |
|                              |           | RNN   | 13.6867            | 11.9887 | 12.78% |
|                              | 87 (90%)  | BLSTM | 3.3549             | 3.0602  | 3.88%  |
|                              |           | LSTM  | 1.8948             | 1.6776  | 2.39%  |
|                              |           | GRU   | 5.0389             | 4.3262  | 5.22%  |
|                              |           | RNN   | 6.7566             | 6.0094  | 5.97%  |

Figura 2.10: Métricas para Avaliação de Diferentes Métodos; Retirada de J. C. Chen et al. (2021)

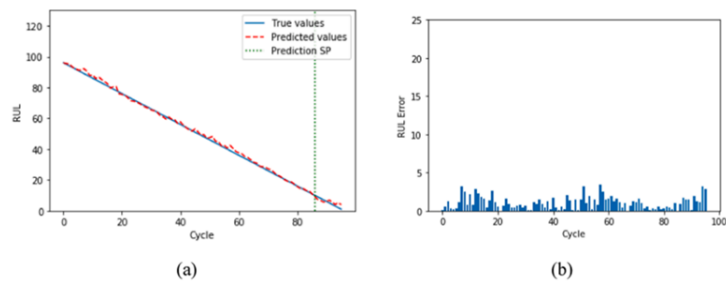


Figura 2.11: a) Previsão de Vida Remanescente Útil para B0018 com SP = 90% b) Erro de Previsão; Retirada de J. C. Chen et al. (2021)



## METODOLOGIA

Neste capítulo, apresenta-se a metodologia adotada neste estudo para a previsão do estado de carga. Inicia-se com a recolha de dados, que envolveu informações de desempenho da bateria, dados ambientais e dados económicos. Em seguida, realiza-se um processo de limpeza e pré-processamento dos dados, seguido pela seleção e extração de características relevantes.

Optou-se por utilizar modelos LSTM para a previsão, considerando tanto abordagens multivariável quanto univariável. O treino dos modelos envolveu ajuste de parâmetros e técnicas de validação. Além disso, foram definidas métricas de avaliação para avaliar o desempenho dos modelos. Por fim, destacam-se as ferramentas utilizadas ao longo do estudo, como Excel, Google Colab, *Plant Monitoring System (PMS)*.

1. **Recolha de dados:** Durante o estágio na Galp, foi realizado um processo exaustivo de recolha de dados para reunir os dados necessários para a previsão do nível da bateria. As fontes de dados foram variadas e incluíram:
  - **Dados de desempenho da bateria:** Os dados do Estado de Carga da bateria e o seu histórico de avarias foram extraídos da base de dados interna da Galp;
  - **Dados ambientais:** Obtidos a partir de centrais meteorológicas e outras fontes externas, incluindo temperatura ambiente, humidade, irradiação solar, chuva e vento. Estes fatores ambientais podem influenciar o desempenho da bateria;
  - **Dados económicos:** Foram recolhidos dados referentes à perda de energia durante os períodos de baixa dos TCUs. Esses dados foram utilizados para obter informações económicas com base nos valores de mercado;
2. **Limpeza e pré-processamento de dados:** Para garantir a qualidade e a adequação dos dados para análise, foi implementado um processo de limpeza e pré-processamento de dados. Foram dados os seguintes passos:
  - Realizou-se uma análise minuciosa dos dados coletados para garantir sua exaustividade, consistência e integridade. Isso envolveu a verificação dos parâmetros-chave, como os níveis de carga da bateria e a humidade relativa, para garantir



que estivessem dentro das faixas operacionais aceitáveis. Além disso, avaliou-se a presença de qualquer repetição de valores em todo o conjunto de dados. Qualquer entrada ausente ou inconsistente foi identificada e devidamente corrigida ou eliminada;

- **Seleção de características:** As características relevantes foram selecionadas com base no seu potencial impacto na previsão do nível da bateria. Para tal, foram tidos em conta os conhecimentos do domínio, as conclusões da revisão da literatura e as técnicas de análise da importância das características, como a análise de correlação;
- **Tratamento de dados em falta:** No âmbito da análise de dados, a questão dos dados em falta foi abordada dependendo da quantidade e natureza dos dados em falta. Para variáveis com poucos valores em falta, onde a perda de dados não afetaria a análise, optou-se pela eliminação desses pontos em falta;
- **Normalização e escalonamento:** Para garantir o tratamento justo de diferentes características e facilitar a convergência do modelo, as características numéricas foram normalizadas ou escalonadas. As técnicas comuns de normalização incluem o escalonamento mínimo-máximo;
- **Pré-processamento de séries temporais:** : Uma vez que os dados de desempenho da bateria apresentam frequentemente dependências temporais, foram utilizadas técnicas de pré-processamento adequadas. Isto envolveu a conversão dos dados de timestamp em representações sinusoidais (seno e cosseno) para melhor capturar e representar as informações temporais contidas nos dados. A razão para essa transformação é que os dados de *timestamp*, por si só, não são adequados para a análise de séries temporais, pois não refletem diretamente a natureza cíclica e periódica das variações temporais. Portanto, a conversão em seno e cosseno permite representar de forma mais eficaz as flutuações temporais, garantindo intervalos de tempo consistentes para análise subsequente.

3. **Desafios e soluções:** Durante o processo de limpeza dos dados, foram encontrados vários desafios que foram efetivamente resolvidos. Esses desafios incluíram:

- **Formatos de dados inconsistentes:** Os dados recolhidos de diferentes fontes tinham formatos e estruturas diferentes. Foram utilizadas técnicas de normalização, como a conversão ou fusão de formatos de dados, para garantir a uniformidade;
- **Erros na introdução de dados:** Algumas entradas de dados continham erros tipográficos ou inconsistências. A inspeção manual e a validação cruzada com outras fontes ajudaram a identificar e a retificar esses erros;
- **Volume de dados:** : Devido a restrições de poder computacional, os dados foram simplificados para torná-los mais ágeis e menos exigentes em termos

---

de recursos. O grande volume de dados coletados exigiu eficiência no armazenamento e processamento, o que foi alcançado com a utilização de sistemas escaláveis e otimização dos recursos computacionais;

4. **Seleção e extração de características relevantes:** Após o processo de limpeza dos dados, foram selecionadas e extraídas características relevantes do conjunto de dados limpo. Isto envolveu uma combinação de conhecimentos do domínio e técnicas de análise de dados como correlações entre os parâmetros disponíveis. As características selecionadas incluíam:
  - **Características de desempenho da bateria:** O SOC considerado um indicador-chave do comportamento e desempenho da bateria;
  - **Dados ambientais:** Obtidos a partir de centrais meteorológicas e outras fontes externas, incluindo temperatura ambiente, humidade e irradiação solar, chuva, vento. Estes fatores ambientais podem influenciar o desempenho da bateria;
5. **Escolha de um modelo:** Foram selecionados modelos LSTM ao fim de um estudo extenso que teve como base a revisão bibliográfica, devido à sua capacidade de captar eficazmente dependências e padrões temporais em dados de séries temporais. Têm mostrado resultados promissores em várias tarefas de previsão de séries temporais, o que as torna adequadas para prever os valores da bateria ao longo do tempo.
6. **Visão geral dos modelos LSTM multivariável e LSTM univariável:** Os modelos LSTM multivariável e LSTM univariável foram considerados para a previsão do nível da bateria ao fim de um estudo e de uma reflexão sobre os modelos LSTM.
  - **Modelo LSTM multivariável:** O modelo LSTM multivariável utiliza várias características de entrada, incluindo dados de desempenho da bateria, dados ambientais, para prever o valor da bateria. Permite que o modelo capte as interações e dependências complexas entre diferentes variáveis, aumentando a exatidão da previsão;
  - **Modelo LSTM Univariável:** O modelo LSTM univariável, por outro lado, concentra-se na previsão do nível da bateria com base apenas nos dados de desempenho da bateria. Não tem em conta a influência de fatores externos, como as variáveis ambientais. Este modelo simplifica a tarefa de previsão, mas pode não captar todos os potenciais fatores de influência;
7. **Processo de treino:** Para ajustar parâmetros e realizar a validação, o processo de treinamento dos modelos LSTM incluiu as seguintes etapas:
  - **Particionamento de dados:** Foi criada uma base de dados dedicada ao treino e à validação, na qual os dados de 20 baterias avariadas foram utilizados para treino, e os dados de 1 bateria avariada foram reservados para validação. Para o conjunto de teste, 11 baterias funcionais e 3 baterias avariadas foram utilizadas;

- **Conceção da arquitetura do modelo:** A arquitetura do modelo LSTM foi cuidadosamente concebida com base em experimentação e evidências empíricas. Nesse processo, foram especificados o número de camadas LSTM, unidades ocultas, taxa de dropout, camadas de normalização por lotes, a função de ativação da camada Dense e o número de unidades na camada Dense. Cada uma dessas escolhas foi criteriosamente considerada para otimizar o desempenho do modelo;
  - **Inicialização e otimização:** Para otimizar o desempenho dos modelos LSTM, o ajuste dos hiperparâmetros foi efetuado utilizando um otimizador Bayesiano. Um otimizador bayesiano recolhe amostras e avalia várias combinações de hiperparâmetros dentro de um intervalo definido. Utiliza um modelo probabilístico para orientar a pesquisa em direção a regiões promissoras do espaço de hiperparâmetros. Os hiperparâmetros que foram ajustados podem incluir a taxa de aprendizagem, o número de camadas LSTM, as unidades ocultas, as taxas de desistência, os comprimentos de sequência ou quaisquer outros hiperparâmetros relevantes específicos dos modelos LSTM;
  - **Treino e validação:** Durante o treinamento dos modelos LSTM, aplicamos a retro-propagação no tempo para ajustar os pesos à medida que a rede neural aprende com os dados do conjunto de treinamento. Ao mesmo tempo, utilizamos o conjunto de validação para avaliar o desempenho do modelo e evitar possíveis problemas de overfitting. Empregamos técnicas de paragem antecipada que nos permitiram interromper o treinamento se o modelo não apresentasse melhorias significativas no desempenho em relação ao conjunto de validação ou se começasse a piorar.
8. **Modelos alternativos considerados:** Para além dos modelos LSTM, foi criado um modelo de base utilizando regressão linear para comparar o seu desempenho com os modelos LSTM. O modelo de regressão linear foi incluído para fornecer uma referência simples e interpretável para a previsão do nível da bateria.

**Modelo de regressão linear:** A regressão linear é uma técnica de modelação estatística amplamente utilizada que pressupõe uma relação linear entre as variáveis independentes e a variável dependente. No contexto da previsão do nível da bateria, o modelo de regressão linear visava estimar o valor da bateria com base nas características de entrada selecionadas. Razões para o usarmos como comparação:

- **Simplicidade e interpretabilidade:** Os modelos de regressão linear são fáceis de compreender e interpretar devido à sua natureza linear. Fornecem informações sobre a relação entre as variáveis independentes e a variável dependente;
- **Aferição de desempenho:** Ao comparar o desempenho dos modelos LSTM com a linha de base da regressão linear, torna-se possível avaliar a eficácia dos

---

modelos LSTM mais complexos na captura das dependências e complexidades temporais presentes nos dados de desempenho da bateria;

9. **Métricas de avaliação:** Definição das métricas de avaliação: Durante a avaliação dos modelos de IA para a previsão do nível da bateria, foi utilizada uma série de métricas para avaliar o seu desempenho. Estas métricas abrangeram vários aspetos da exatidão da previsão, do ajuste do modelo e da exatidão. Foram utilizadas as seguintes métricas de avaliação:
- **Mean Squared Error (MSE):** O MSE mede a diferença média ao quadrado entre os valores previstos da bateria e os valores reais. Ele fornece uma medida da capacidade do modelo de minimizar os erros gerais de previsão;
  - **Mean Absolute Error (MAE):** O MAE calcula a diferença média absoluta entre os valores previstos da bateria e os valores reais. Representa a magnitude média dos erros, independentemente da sua direção;
  - **Root Mean Squared Error (RMSE):** O RMSE é a raiz quadrada da média das diferenças quadradas entre os valores previstos e os valores reais. É uma métrica popular que fornece uma medida da exatidão da previsão, dando maior peso a erros maiores;
  - **Mean Absolute Percentage Error (MAPE):** O MAPE calcula a diferença percentual média entre os valores previstos e reais. Quantifica a exatidão relativa das previsões e é particularmente útil quando se interpreta o erro em termos percentuais;
  - **Coeficiente de determinação (R-quadrado):** O R-quadrado mede a proporção da variação na variável dependente (SOC) que pode ser explicada pelas variáveis independentes (características de entrada). Oferece uma indicação de quão bem o modelo se ajusta aos dados;
  - **Exatidão:** A exatidão é uma métrica que compara cada ponto previsto com o valor real, considerando-o correto se a diferença for menor que 5% e incorreto se a diferença for maior que 5%. Posteriormente, calcula-se a percentagem de previsões corretas em toda a amostra.
10. **Ferramentas Utilizadas:** As ferramentas utilizadas ao longo deste estudo foram o Excel e o Google Colab (Python), devido ao maior poder computacional do *Graphics Processing Unit (GPU)* e o *PMS* da Galp.

## CASO DE ESTUDO

O Conjunto de interesse neste estudo é denominado por "Alcazar", composto por parques distintos como pode ser observado pela Figura 4.1: Alcázar I (44,99 MWp), Alcázar II (44,96 MWp), Valdivieso (49,88 MWp) e Valdecarro (49,88 MWp). Localizado em Espanha, o parque possui uma potência de 190 MWp em corrente contínua (DC) e 175 MVA em corrente alternada (AC).



Figura 4.1: Parque "Alcazar"190MWp.

Na Figura 4.2 pode ser observado as características de cada parque fotovoltaico. Pode ser observado em maior detalhe o parque na Figura 4.3.

| Alcazar                   |                        |                         |                         |                        |
|---------------------------|------------------------|-------------------------|-------------------------|------------------------|
| Parque                    | Alcázar I ( 44,99 MWp) | Alcázar II ( 44,96 MWp) | Valdivieso ( 49,88 MWp) | Valdecarro (49,88 MWp) |
| Strings                   | 4.396                  | 4.391                   | 4.913                   | 4.913                  |
| Estruturas de suporte     | 1.465                  | 1.463                   | 1.437                   | 1.437                  |
| Configuração              | 2Vx45                  | 2Vx45                   | 2Vx45                   | 2Vx45                  |
| Potência dos módulos (Wp) | 340 Wp                 | 101.280                 | 89.610                  | 99.870                 |
|                           | 345Wp                  | 30.600                  | 42.120                  | 47.520                 |
| Total de Módulos          | 131.880                | 131.730                 | 147.390                 | 147.520                |
| Rotação Max               | ±55°                   | ±55°                    | ±55°                    | ±55°                   |
| TCUs                      | 1.465                  | 1.463                   | 1.437                   | 1.437                  |
| NCUs                      | 12                     | 12                      | 12                      | 12                     |
| Inversores                | 12                     | 12                      | 13                      | 13                     |
| N° de Paineis por String  | 90                     | 90                      | 90                      | 90                     |

Figura 4.2: Descrição das características principais dos parques fotovoltaicos



Figura 4.3: Transformador e Inversor de “Alcazar”.

A instalação de Valdecarro possui uma subestação, para a qual convergem todos os cabos de média tensão. São utilizados 500 metros de cabos subterrâneos para a passagem desses cabos até a subestação como podemos ver pela Figura 4.4. Posteriormente, a energia é convertida para alta tensão e transferida para outra subestação próxima chamada Manchasol, de onde é distribuída para a rede.



Figura 4.4: Subestação de "Alcazar".

O objetivo deste estudo é analisar a variação de carga da bateria. Por isso, foi realizado um estudo das correlações entre as variáveis que mais influenciam o SOC dos rastreadores fotovoltaicos.



## ANÁLISE DOS RESULTADOS

Neste capítulo, será dedicado espaço para a apresentação e análise dos resultados obtidos no estudo. Inicialmente, será realizada uma exploração das correlações entre as variáveis do conjunto de dados. Serão identificadas possíveis relações e dependências entre essas variáveis. Além disso, será realizada a visualização dos dados recolhidos, por meio de gráficos e representações visuais. A criação de uma *baseline* servirá como referência inicial para a avaliação dos resultados dos modelos desenvolvidos.

Os resultados do modelo LSTM multivariável serão apresentados e analisados. Serão discutidas as métricas de desempenho utilizadas, a fim de avaliar a qualidade e a exatidão das previsões geradas pelo modelo.

Em seguida, serão abordados os resultados do modelo LSTM univariável. As métricas de desempenho serão utilizadas para avaliar a qualidade das previsões. Será fornecida uma análise comparativa entre os resultados do modelo LSTM multivariável e univariável.

Por fim, será realizada uma simulação económica baseada nos resultados dos modelos. Essa simulação explorará os impactos financeiros e os benefícios decorrentes da implementação das previsões geradas.

### 5.1 Correlações

O propósito deste estudo consiste em examinar as flutuações na carga das baterias a fim de identificar quais delas apresentam falhas ou avarias. Por isso, foi realizado um estudo das correlações entre as variáveis que mais influenciam o SOC dos rastreadores fotovoltaicos.

Pretende-se compreender a influência de cada uma das variáveis na carga da bateria (*Battery level(%)*), conforme observado na Figura 5.1. O gráfico possui uma escala que varia de -1 a 1, em que valores negativos indicam que, quando uma variável aumenta, a outra tende a diminuir, enquanto valores positivos descrevem uma variação positiva, indicando que quanto maior o valor, maior a influência positiva da outra variável. Portanto, quando uma variável aumenta, a outra também aumenta.

As variáveis exercem influência no nível de bateria de várias maneiras. Por exemplo, a **humidade relativa** (*Moisture (%)*) contribui não apenas para a degradação da condição

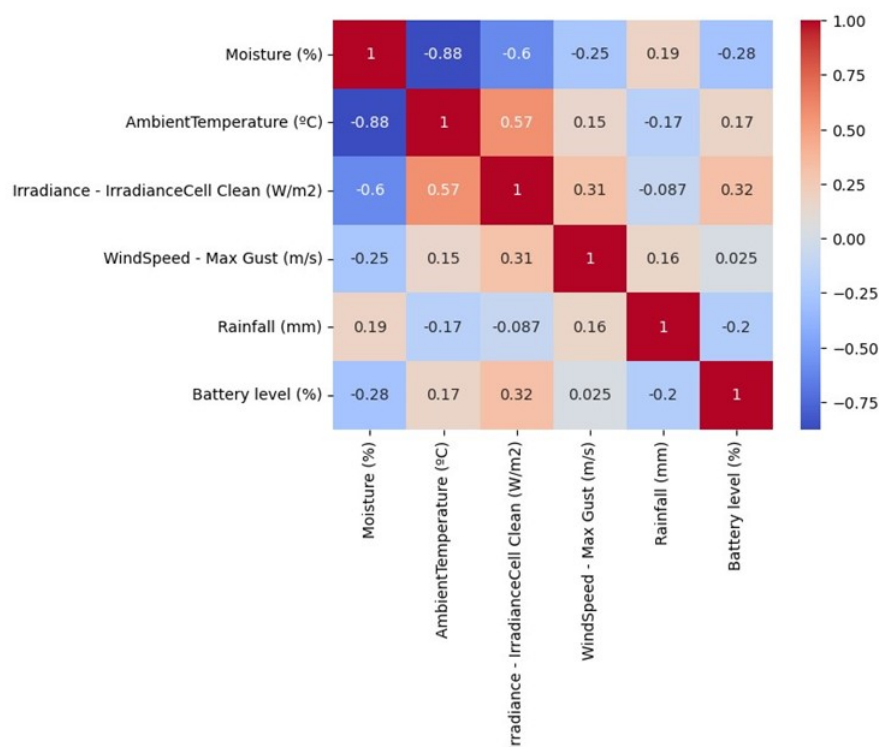


Figura 5.1: Correlação das variáveis com maior influência no nível de bateria.

da bateria, mas também para a redução da quantidade de fótons que alcançam os painéis fotovoltaicos. Portanto, ela possui um valor consideravelmente negativo.

A **temperatura ambiente** (*Ambient Temperature (C)*) tem diversos impactos no SOC. Em altas temperaturas, seu impacto pode ser considerado negativo devido à perda de eficiência causada nos painéis fotovoltaicos. Porém, o aspecto mais predominante é o positivo, pois a temperatura e a irradiância estão fortemente relacionadas. Assim, um aumento na irradiância tende a aumentar a temperatura ambiente, que, por sua vez, tende a aumentar o nível de carga da bateria.

A **irradiância** utilizada neste estudo foi a "*Irradiance - IrradianceCell Clean (W/m2)*". Optou-se por essa medida devido às falhas de informação encontradas nos valores do piranômetro. Essa informação é obtida a partir de dois pequenos módulos adjacentes, nos quais um deles é regularmente limpo, enquanto o outro só é limpo durante a manutenção geral dos módulos no parque. A utilização da "*IrradianceCell Clean*" tem como propósito a comparação com a "*IrradianceCell Soiled*", a fim de avaliar o nível de sujeira nos painéis e identificar as perdas decorrentes do efeito de sujeira. A irradiância mencionada é submetida a limpezas regulares e oferece dados mais confiáveis e consistentes devido à manutenção constante realizada.

A **velocidade máxima do vento** (*Windspeed - Max Gust*) também possui vários impactos. Por exemplo, possui uma relação negativa com o SOC, uma vez que altas velocidades de vento exigem uma grande quantidade de energia dos seguidores fotovoltaicos para posicionar os painéis em modo de defesa. No entanto, possui uma relação bastante



positiva e forte com a irradiância, e essas duas influências contrabalançam-se. Pode dar a impressão de que o seu impacto é relativamente pequeno, mas, na realidade, este tem um impacto direto negativo considerável, uma vez que a necessidade de ajustar a posição de defesa consome energia da bateria. Além disso, verifica-se um impacto indireto positivo substancial, uma vez que o vento aumenta a irradiância, resultando numa maior alimentação de energia para a bateria. No entanto, o impacto que prevalece é um pequeno impacto positivo.

A **chuva** (*Rainfall (mm)*) em geral leva à ausência de radiação solar, o que tende a diminuir o nível de carga da bateria. Além disso, a chuva geralmente está associada a ventos fortes e altos níveis de humidade, o que torna negativa sua relação com o SOC. No entanto, a sua influência no SOC ocorre principalmente pela forma como é afetada pelas outras variáveis.

## 5.2 Visualização dos dados e criação de uma *Baseline*

As informações relativas a cada uma das variáveis referidas anteriormente podem ser visualizadas na Figura 5.2, que apresenta os valores utilizados para a construção da base de dados. É importante salientar que o nível de chuva não foi incluído devido à falta de dados disponíveis. No entanto, foram utilizados diferentes momentos no tempo e dados de diferentes campos fotovoltaicos para criar essa base de dados. O intervalo de tempo utilizado foi de horas. Por meio dessa figura, é possível obter uma representação visual das diferentes variáveis e seus respectivos valores ao longo do tempo, permitindo observar as variações e padrões ao longo do período analisado.

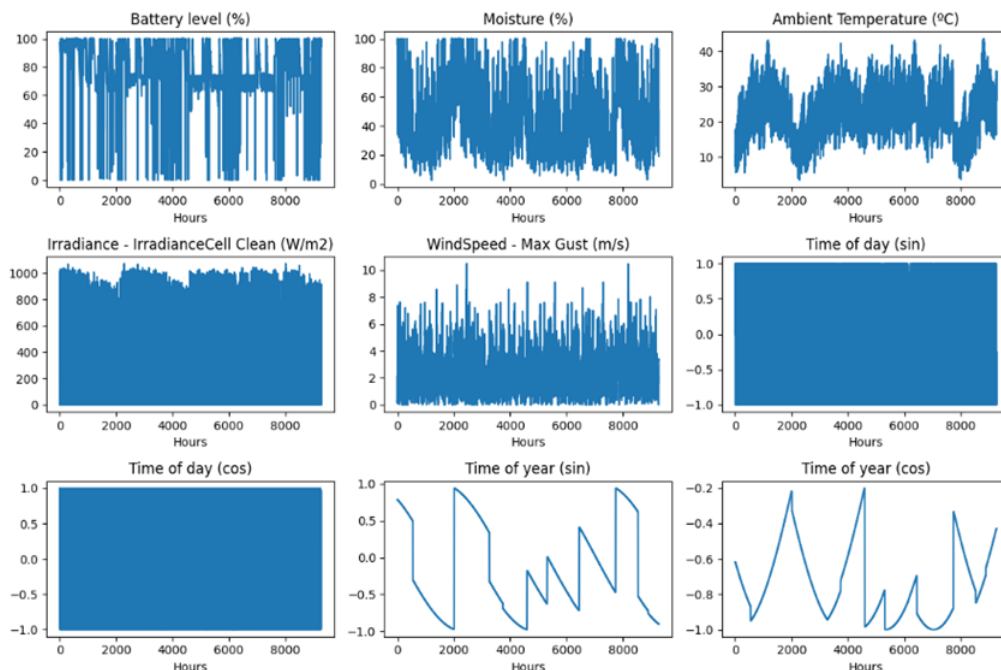


Figura 5.2: Gráficos das Variáveis Usadas para a Criação do Modelo LSTM Multivariada.

Para as previsões foi criado uma base de dados com estas variáveis. Primeiro foi criada uma baseline em que se utilizou um algoritmo de *machine learning*, uma regressão linear *vanilla*, para servir de comparação com os modelos de *deep learning* utilizados e estes são os seus resultados:

Os resultados da *baseline* utilizando um algoritmo de *machine learning*, como a regressão linear *vanilla*, foram obtidos para servir de comparação com os modelos de *deep learning* utilizados. Podem ser observados os resultados obtidos na Figura 5.3:

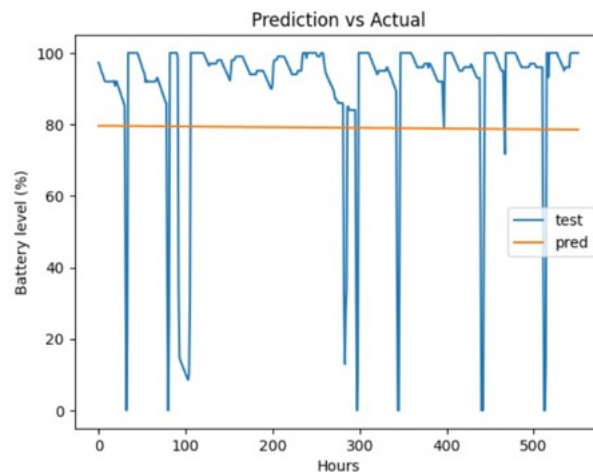


Figura 5.3: Regressão Linear - Previsão vs Valores reais.

A decisão de não utilizar a regressão linear na comparação com os modelos LSTM baseou-se na análise das Figuras 5.5, 5.6 e na comparação com a Figura 5.4

Observamos que a regressão linear apresenta limitações na deteção de anomalias no SOC, as quais desencadeiam o alarme de falha na bateria. Embora a regressão linear consiga seguir o padrão geral ao acompanhar a média dos valores de SOC, não é suficientemente sensível para identificar as variações específicas que indicam problemas nas baterias. Consequentemente, concluímos que a regressão linear não seria útil para esta aplicação, pois não fornece resultados satisfatórios na deteção de falhas nas baterias. Portanto, não a utilizaremos nos testes subsequentes. As métricas desse modelo podem ser vistas na Figura 5.5 para fins de comparação com os testes futuros.

| Métricas   |        |
|------------|--------|
| MSE:       | 587,56 |
| MAE:       | 16,89  |
| RMSE:      | 24,24  |
| MAPE:      | inf    |
| R2:        | 0,013  |
| Precision: | 18.70% |

Figura 5.4: Tabela de métricas obtidas pela regressão linear.

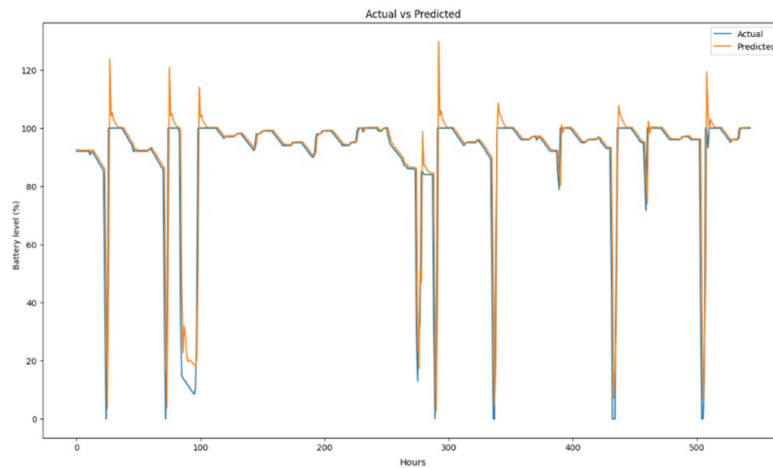


Figura 5.5: LSTM univariável - Previsão vs Valores reais.

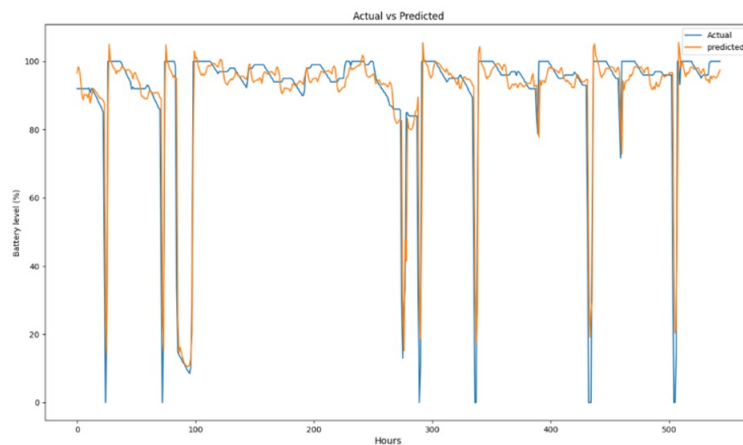


Figura 5.6: LSTM univariável - Previsão vs Valores reais.

### 5.3 Resultados do Modelo LSTM Multivariável

Após a obtenção das informações sobre as correlações entre as variáveis, foi desenvolvido um modelo LSTM multivariável que utiliza todas essas variáveis como entradas e tem como saída o SOC. Com o objetivo de otimizar os hiperparâmetros desse modelo, foi utilizado o *BayesianOptimizer*.

Por meio do *BayesianOptimizer* foram realizadas várias iterações ajustando os hiperparâmetros como o número de neurônios, os valores de *dropout*, os valores de *dense*, as funções de ativação e o *learning rate*. Dessa forma, foram feitas diversas tentativas até encontrar a configuração que melhor se adequasse ao problema em questão.

Para testar esse modelo, foram analisados 11 momentos em que ocorreram anomalias nos seguidores dos 4 campos fotovoltaicos, resultando no disparo do alarme de falha de bateria. No entanto, as baterias não precisaram de ser substituídas em nenhum desses casos. Além disso, foram analisadas 3 baterias que, após o alarme, precisaram ser substituídas, envolvendo a substituição de um TCU.

A seguir, apresentam-se os resultados da previsão do modelo para uma bateria no fim

da sua vida útil em comparação com os valores reais, é relevante notar que esta previsão é feita com base nos últimos 8 valores e prevê apenas um valor no futuro.

Estes podem ser observados pelas Figuras 5.7 a 5.20. É importante notar que o intervalo de tempo escolhido para os testes foi em horas. Os resultados fornecerão uma avaliação do desempenho do modelo LSTM multivariável na previsão do SOC:

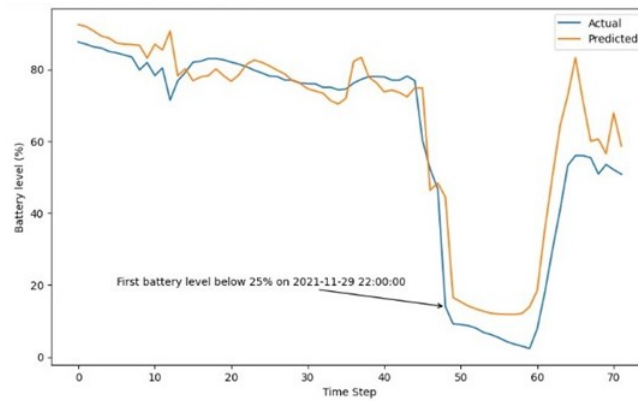


Figura 5.7: Previsão vs Real Bateria Funcional 1.

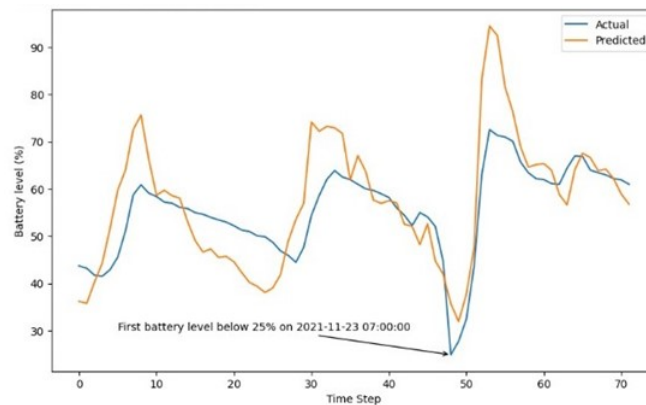


Figura 5.8: Previsão vs Real Bateria Funcional 2.

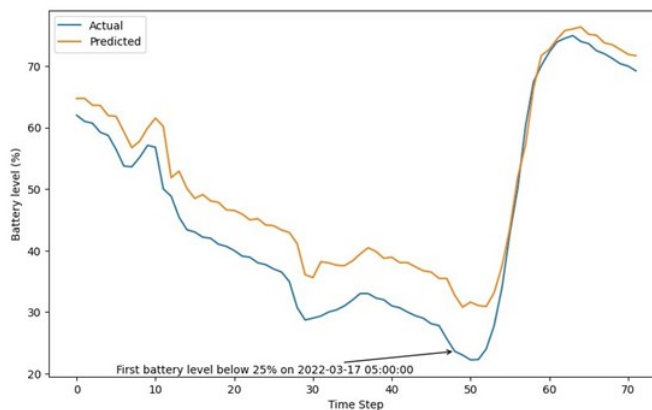


Figura 5.9: Previsão vs Real Bateria Funcional 3.

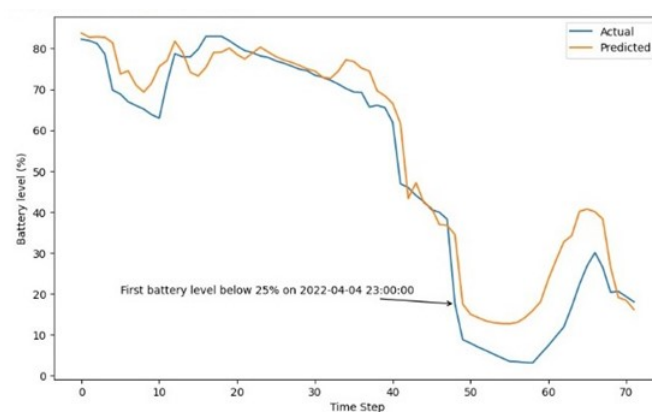


Figura 5.10: Previsão vs Real Bateria Funcional 4.

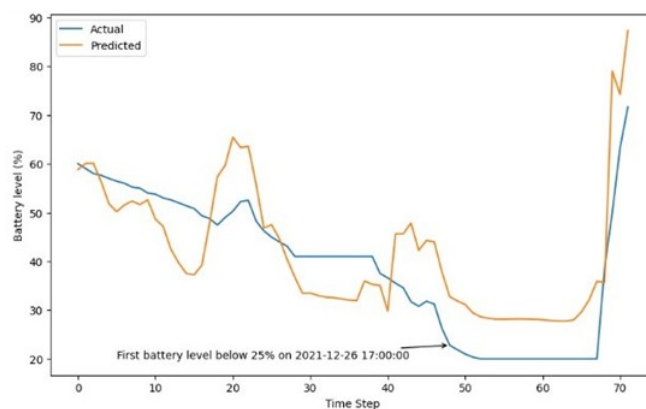


Figura 5.11: Previsão vs Real Bateria Funcional 5.

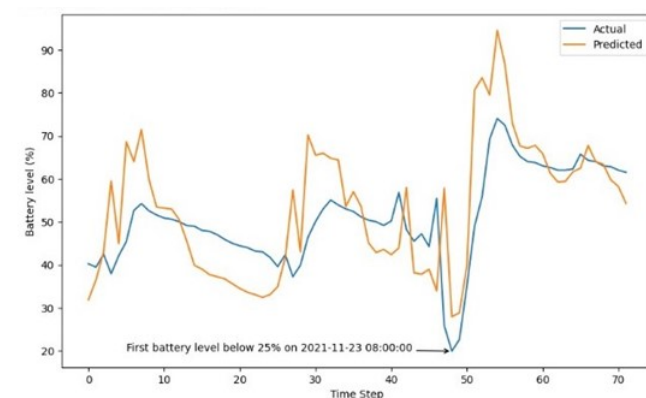


Figura 5.12: Previsão vs Real Bateria Funcional 6.

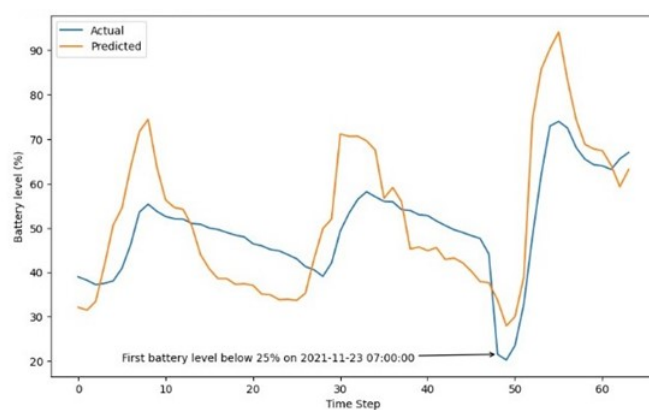


Figura 5.13: Previsão vs Real Bateria Funcional 7.

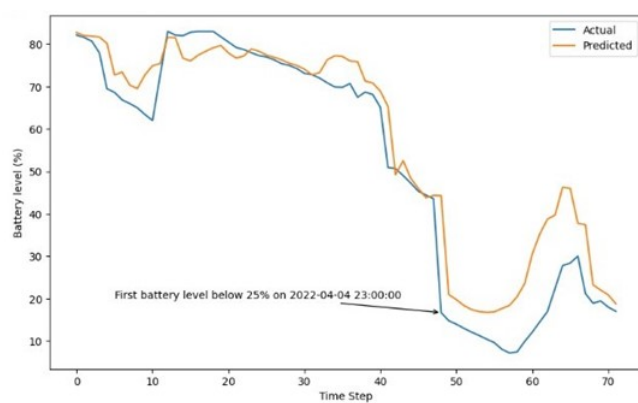


Figura 5.14: Previsão vs Real Bateria Funcional 8.

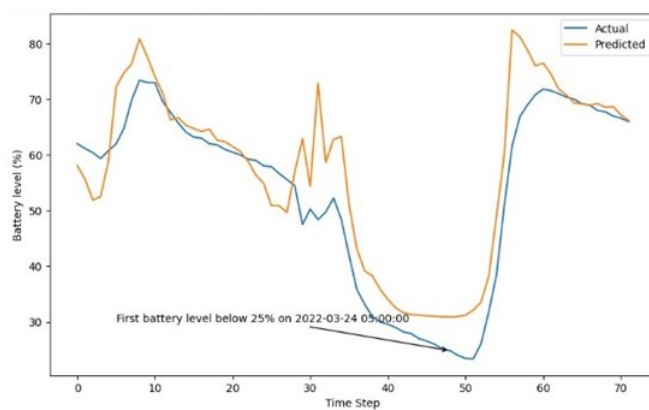


Figura 5.15: Previsão vs Real Bateria Funcional 9.

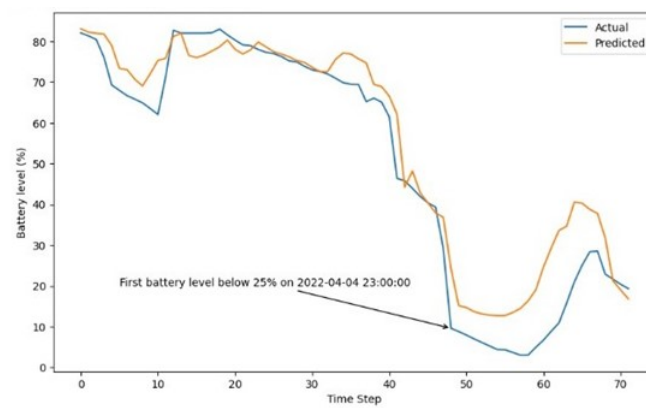


Figura 5.16: Previsão vs Real Bateria Funcional 10.

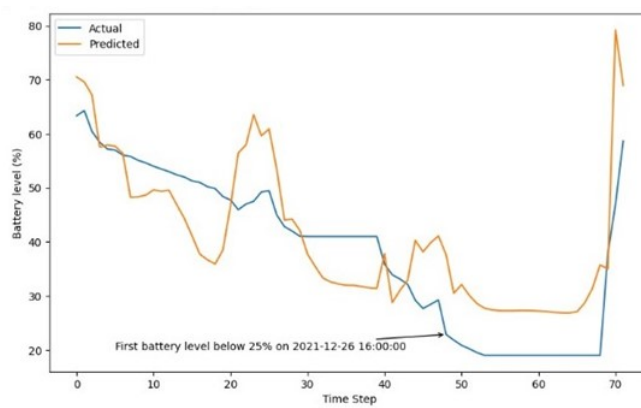


Figura 5.17: Previsão vs Real Bateria Funcional 11.

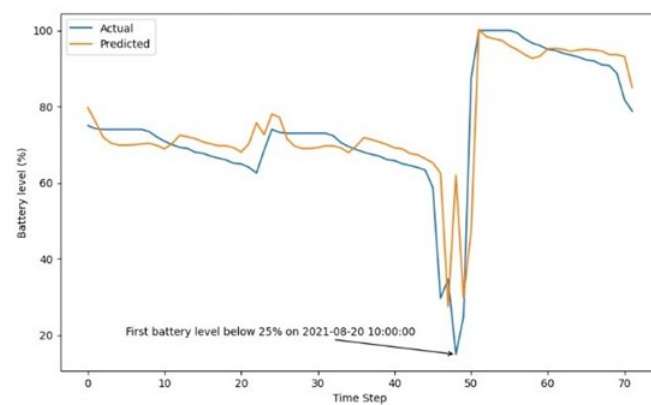


Figura 5.18: Previsão vs Real Bateria Avariada A.

### 5.3. RESULTADOS DO MODELO LSTM MULTIVARIÁVEL

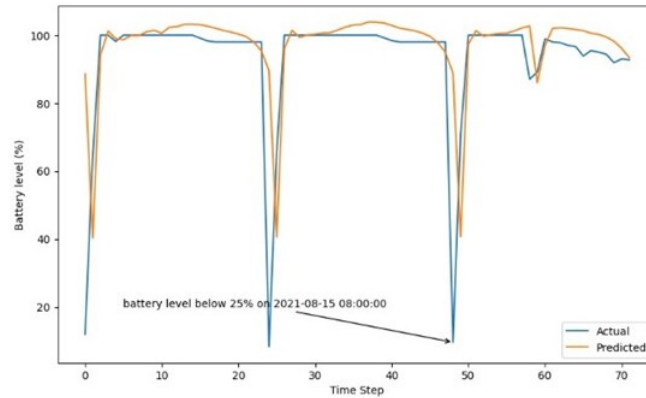


Figura 5.19: Previsão vs Real Bateria Avariada B.

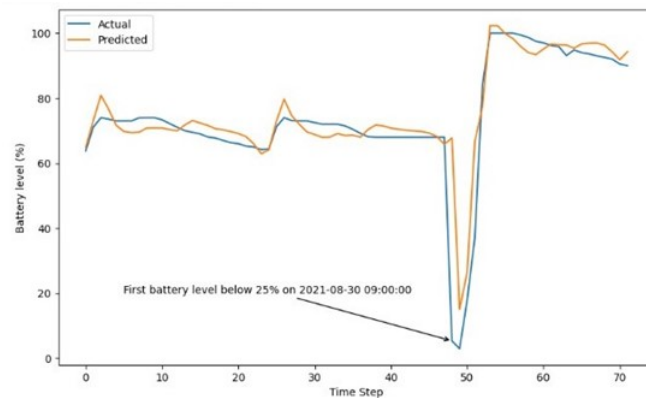


Figura 5.20: Previsão vs Real Bateria Avariada C.

Foi criada a Figura 5.21 para se fazer a comparação com as várias métricas:

| Métricas              | Baterias Funcionais |       |       |       |       |        |        |       |       |       |       | Baterias Avariadas |        |       |
|-----------------------|---------------------|-------|-------|-------|-------|--------|--------|-------|-------|-------|-------|--------------------|--------|-------|
|                       | 1                   | 2     | 3     | 4     | 5     | 6      | 7      | 8     | 9     | 10    | 11    | A                  | B      | C     |
| MSE                   | 82,72               | 68,97 | 56,36 | 64,26 | 88,24 | 127,08 | 115,03 | 73,48 | 50,66 | 69,52 | 88,04 | 83,75              | 301,48 | 77,70 |
| MAE                   | 6,54                | 6,43  | 6,19  | 6,06  | 8,19  | 8,52   | 9,13   | 6,08  | 5,33  | 6,23  | 8,07  | 5,00               | 6,92   | 4,06  |
| RMSE                  | 9,10                | 8,30  | 7,51  | 8,02  | 9,39  | 11,27  | 10,73  | 8,57  | 7,12  | 8,34  | 9,38  | 9,15               | 17,36  | 8,81  |
| MAPE (%)              | 39,10               | 12,09 | 15,67 | 49,52 | 25,48 | 18,53  | 19,26  | 29,93 | 12,20 | 53,19 | 26,26 | 11,44              | 38,07  | 26,99 |
| R2                    | 0,90                | 0,22  | 0,80  | 0,92  | 0,56  | -0,10  | 0,06   | 0,90  | 0,81  | 0,92  | 0,58  | 0,70               | 0,13   | 0,77  |
| Exatidão (%)          | 54,17               | 47,22 | 45,83 | 52,78 | 22,22 | 40,28  | 23,44  | 54,17 | 52,78 | 48,61 | 22,22 | 86,11              | 81,94  | 90,28 |
| Média de Exatidão (%) | 42,16               |       |       |       |       |        |        |       |       |       |       | 86,11              |        |       |

Figura 5.21: Métricas de Previsão VS Real do Modelo LSTM Multivariável.

Pela análise da Figura 5.21, é possível observar que os valores de exatidão são consideravelmente maiores nas baterias que estavam realmente estragadas e precisaram ser substituídas, com uma média de exatidão de 86,11%. Por outro lado, as baterias funcionais, que tiveram apenas falhas temporárias devido a anomalias ou a condições climáticas desfavoráveis, são aquelas afetadas por semanas de baixa exposição solar, ou seja, períodos com pouca incidência de luz solar, apresentaram uma média de exatidão de 42,16%. Essa diferença na exatidão permite distinguir entre as baterias que acionaram o alarme de



falha de bateria e precisaram ser substituídas e aquelas que tiveram apenas uma falha temporária.

No entanto, as demais métricas observadas não fornecem uma distinção clara entre as baterias que estavam avariadas e aquelas que tiveram apenas uma anomalia que ativou o alarme. Isso indica que outras métricas não são tão eficazes na diferenciação dessas situações. Portanto, a **exatidão** destaca-se como a métrica mais relevante para essa distinção.

Estes resultados destacam a importância da exatidão como uma medida confiável para avaliar o desempenho do modelo na detecção de baterias danificadas que precisam ser substituídas, possibilitando a diferenciação entre falhas temporárias e falhas que requerem intervenção.

## 5.4 Resultados do Modelo LSTM Univariável

Para este estudo, além da regressão linear vanilla e do modelo LSTM multivariável, também foi utilizado o modelo LSTM multivariável de múltiplos passos, que não apresentou resultados tão favoráveis. Além desses modelos, também foi empregue o modelo LSTM univariável de múltiplos passos, que também não obteve bons resultados. No entanto, o modelo LSTM univariável demonstrou um desempenho satisfatório, conforme pode ser observado nas Figuras 5.22 a 5.35.

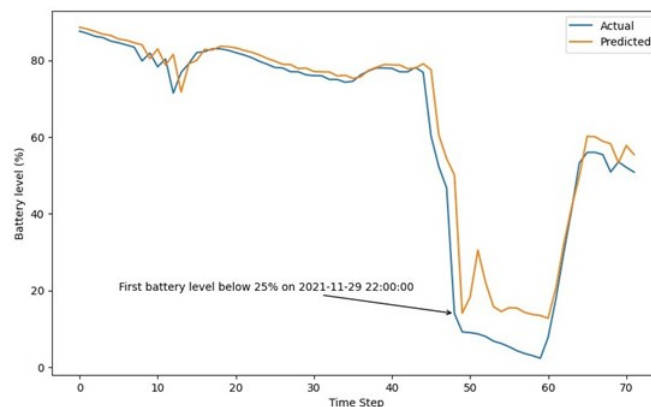


Figura 5.22: Previsão vs Real Bateria Funcional 1.

#### 5.4. RESULTADOS DO MODELO LSTM UNIVARIÁVEL

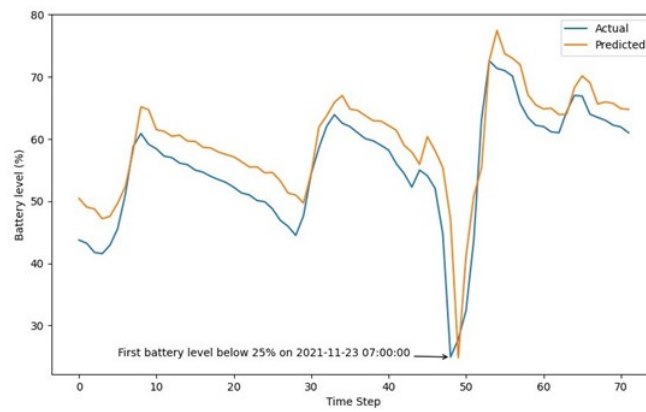


Figura 5.23: Previsão vs Real Bateria Funcional 2.

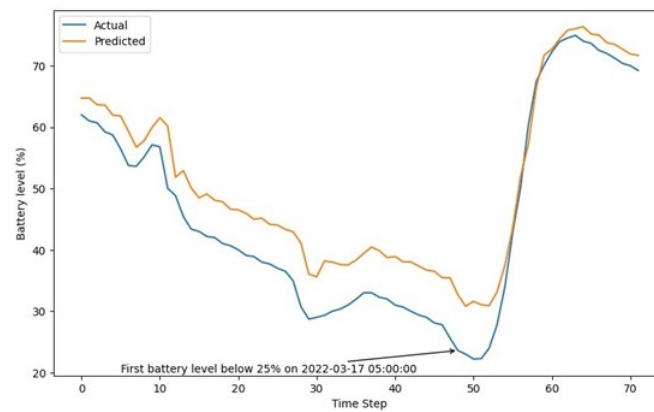


Figura 5.24: Previsão vs Real Bateria Funcional 3.

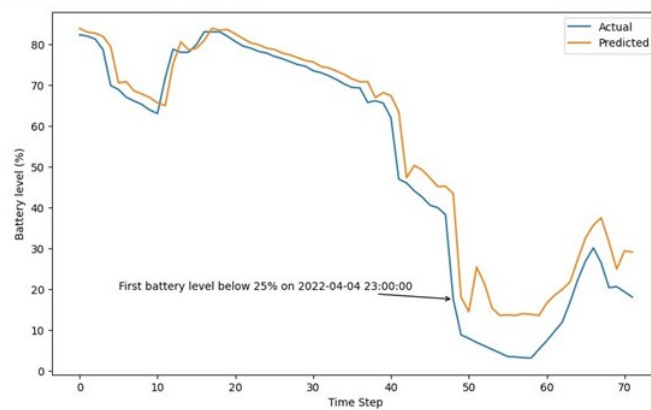


Figura 5.25: Previsão vs Real Bateria Funcional 4.

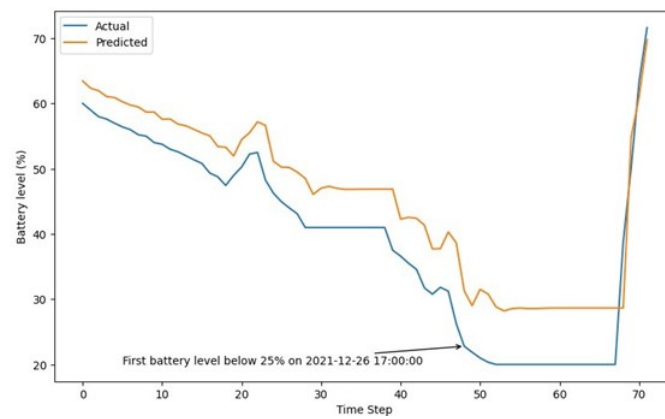


Figura 5.26: Previsão vs Real Bateria Funcional 5.

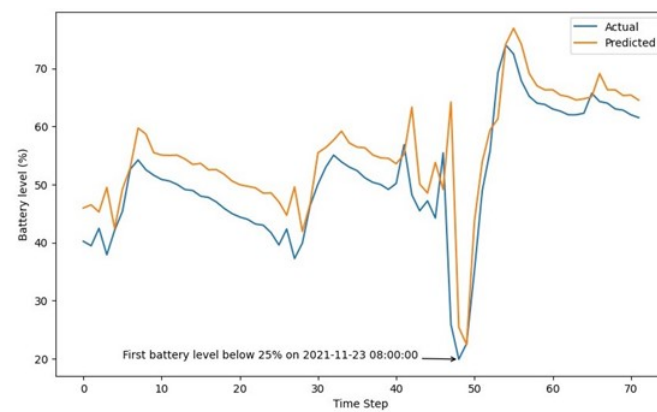


Figura 5.27: Previsão vs Real Bateria Funcional 6.

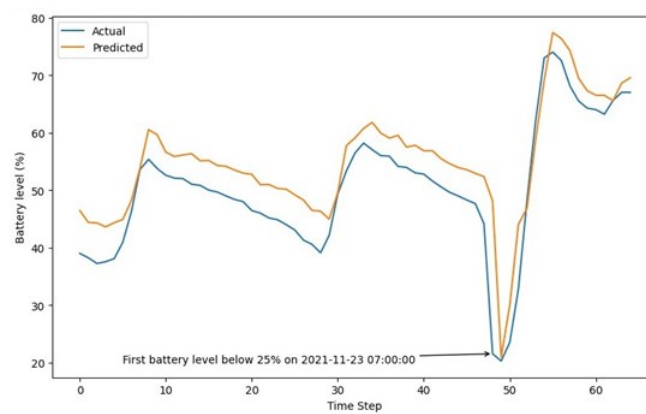


Figura 5.28: Previsão vs Real Bateria Funcional 7.

#### 5.4. RESULTADOS DO MODELO LSTM UNIVARIÁVEL

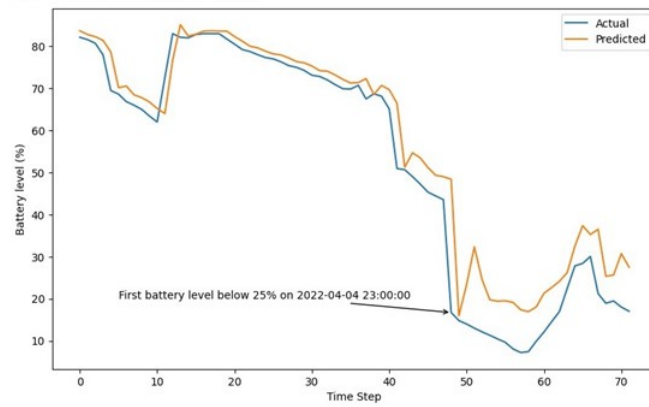


Figura 5.29: Previsão vs Real Bateria Funcional 8.

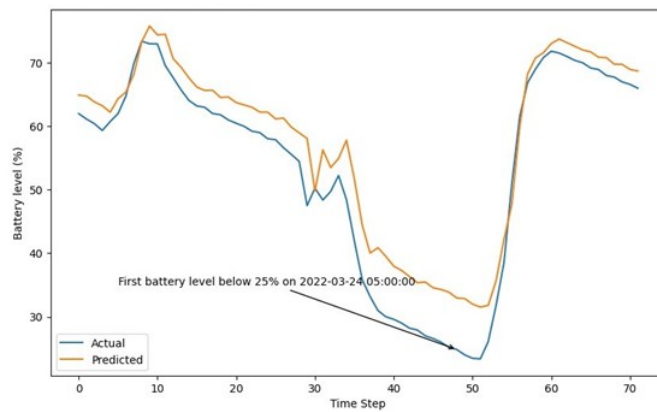


Figura 5.30: Previsão vs Real Bateria Funcional 9.

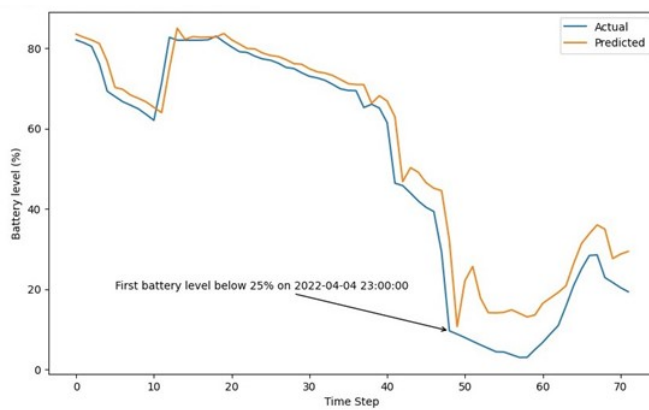


Figura 5.31: Previsão vs Real Bateria Funcional 10.

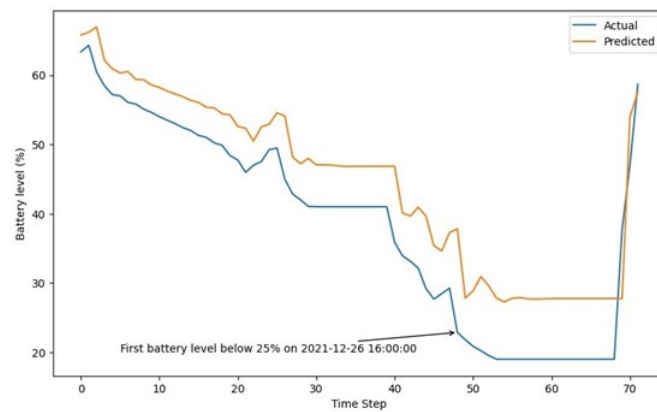


Figura 5.32: Previsão vs Real Bateria Funcional 11.

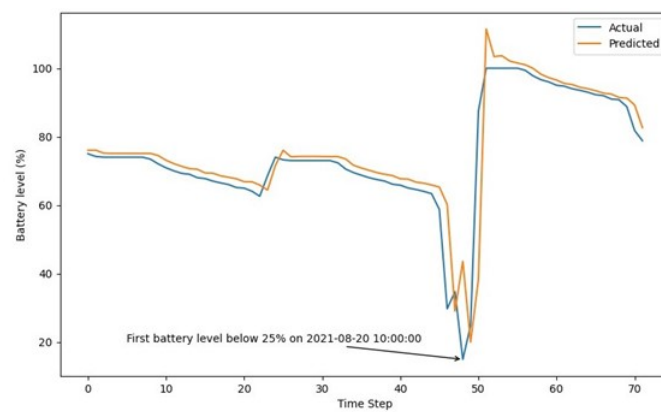


Figura 5.33: Previsão vs Real Bateria Avariada A.

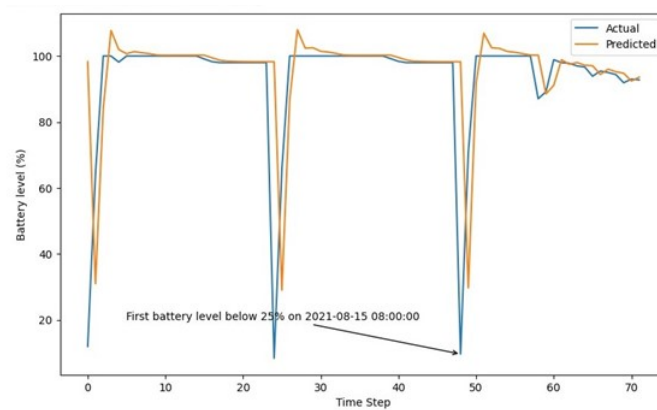


Figura 5.34: Previsão vs Real Bateria Avariada B.

#### 5.4. RESULTADOS DO MODELO LSTM UNIVARIÁVEL

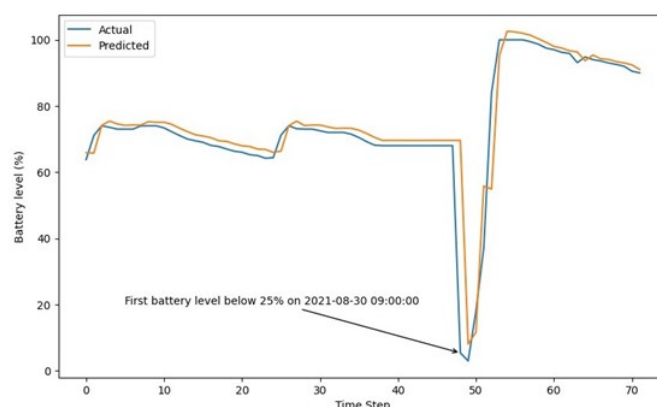


Figura 5.35: Previsão vs Real Bateria Avariada C.

Na Figura 5.36, apresentam-se as métricas relacionadas com o modelo LSTM univariável. Essas métricas são utilizadas para avaliar o desempenho e a exatidão do modelo na previsão do SOC com base em apenas uma variável de entrada. Ao analisar as métricas, é possível ter uma visão mais clara do desempenho do modelo e da sua capacidade de fazer previsões exatas.

| Métricas              | Baterias Funcionais |       |       |       |       |       |       |       |       |       |       | Baterias Avariadas |        |       |
|-----------------------|---------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------------------|--------|-------|
|                       | 1                   | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | 11    | A                  | B      | C     |
| MSE                   | 51,05               | 27,70 | 36,45 | 51,35 | 46,19 | 47,97 | 34,88 | 55,04 | 27,34 | 51,63 | 48,66 | 65,59              | 395,70 | 78,59 |
| MAE                   | 4,22                | 4,41  | 5,40  | 5,33  | 6,40  | 5,03  | 4,87  | 5,26  | 4,41  | 5,41  | 6,53  | 3,69               | 7,04   | 3,38  |
| RMSE                  | 7,14                | 5,26  | 6,04  | 7,17  | 6,80  | 6,93  | 5,91  | 7,42  | 5,23  | 7,19  | 6,98  | 8,10               | 19,89  | 8,87  |
| MAPE (%)              | 41,69               | 9,12  | 15,59 | 48,35 | 21,46 | 11,57 | 11,49 | 28,05 | 11,50 | 50,02 | 22,82 | 8,10               | 41,91  | 22,97 |
| R2                    | 0,94                | 0,69  | 0,87  | 0,94  | 0,77  | 0,58  | 0,72  | 0,93  | 0,90  | 0,94  | 0,77  | 0,77               | -0,14  | 0,77  |
| Exatidão (%)          | 75,00               | 68,06 | 38,89 | 56,94 | 34,72 | 58,33 | 56,92 | 59,72 | 70,83 | 51,39 | 29,17 | 90,28              | 80,56  | 91,67 |
| Média de Exatidão (%) | 54,54               |       |       |       |       |       |       |       |       |       |       | 87,50              |        |       |

Figura 5.36: Métricas de Previsão VS Real do Modelo LSTM Univariável.

É evidente, ao observar os resultados apresentados, que a exatidão das baterias que estavam realmente avariadas é significativamente superior, alcançando 87,50%. Por outro lado, as baterias funcionais apresentaram uma exatidão de apenas 54,54%. Essa diferença sugere que o modelo tem a capacidade de detectar, a partir do primeiro alarme de falha de bateria, se há necessidade de substituição da mesma ou não. Essa capacidade pode resultar em economia de recursos e possibilitar uma gestão mais eficiente dos mesmos. Uma vez que a rápida identificação de baterias defeituosas permite direcionar os recursos de substituição apenas para os casos necessários, evitando gastos desnecessários com baterias que ainda estão em boas condições.

Pode-se argumentar que o modelo LSTM multivariável apresenta uma melhor capacidade na detecção de baterias avariadas em comparação ao modelo LSTM univariável. Isso deve-se à diferença nas variações das médias de exatidão entre os dois modelos. No modelo LSTM univariável, a variação das médias de exatidão é de 32,96%, enquanto no modelo LSTM multivariável, essa variação é de 43,95%. Essa diferença indica que o modelo LSTM multivariável possui uma capacidade mais robusta de identificar baterias

avariadas com maior exatidão, tornando-se uma opção mais eficiente para essa finalidade.

## 5.5 Simulação Económica

| Mês 2022  | Valor Médio Mensal (€/MWh) | Perda (MWh) | Perda (€) |
|-----------|----------------------------|-------------|-----------|
| Janeiro   | 201,79                     | 1,225       | 247 €     |
| Fevereiro | 200,22                     | 8,41        | 1 685 €   |
| Março     | 283,39                     | 3,45        | 978 €     |
| Abril     | 191,52                     | 11,30       | 2 165 €   |
| Maio      | 187,13                     | 1,71        | 319 €     |
| Junho     | 169,63                     | 0,87        | 147 €     |
| Julho     | 142,66                     | 1,37        | 195 €     |
| Agosto    | 154,90                     | 1,10        | 170 €     |
| Setembro  | 141,07                     | 1,53        | 216 €     |
| Outubro   | 127,22                     | 0,77        | 98 €      |
| Novembro  | 115,56                     | 16,11       | 1 861 €   |
| Dezembro  | 99,96                      | 63,01       | 6 299 €   |
| Total     |                            | 110,85      | 14 380 €  |

Figura 5.37: Simulação Económica das perdas energéticas por falhas dos seguidores fotovoltaicos no “Alcazar 1”.

Na Figura 5.37, foram utilizados dados provenientes da fonte (**analysis2023**) "esios rede elétrica" para o valor médio mensal do MWh. Foi realizada uma simulação das perdas decorrentes de atrasos no arranjo dos seguidores fotovoltaicos, determinando-se a quantidade de energia perdida. Os valores das perdas foram cuidadosamente calculados por engenheiros da Galp. Para esse cálculo, utilizaram-se dados meteorológicos reais e estimou-se a produção como se os seguidores estivessem na posição ideal. A diferença entre essa estimativa e a produção real foi calculada, sendo que apenas se consideraram os casos em que coincidia com a ocorrência de um alarme de falha de bateria e não havia energia em excesso no mesmo período.

Adicionalmente, a implementação deste programa poderia resultar numa redução da dependência de mão de obra, o que permitiria uma diminuição dos custos de pessoal afeto à manutenção. Logo, os 14.379 euros apurados representam uma estimativa conservadora do valor económico que poderia estar a ser perdido num único parque fotovoltaico com a capacidade de 50 MWp como “Alcazar1”.

## CONCLUSÃO

### 6.1 Conclusão

Em conclusão, este estudo abordou a determinação do estado das baterias dos seguidores fotovoltaicos com o auxílio da inteligência artificial, conseguindo estabelecer uma distinção nítida de exatidão entre as baterias em funcionamento e as avariadas, utilizando a métrica de exatidão. Ao rever o estado atual da manutenção preditiva em centrais fotovoltaicas, ficou evidente que a manutenção desempenha um papel crítico, trazendo benefícios como maior fiabilidade, redução de custos e aumento na produção de energia. Embora a manutenção preventiva seja amplamente empregue atualmente, a manutenção preditiva surge como o futuro neste campo devido aos seus benefícios significativos.

A análise das correlações revelou que a humidade relativa, a temperatura ambiente, a radiação e a velocidade máxima do vento apresentam relações significativas com o nível de carga da bateria dos rastreadores fotovoltaicos. Essas informações são essenciais para entender os fatores que influenciam o estado dos equipamentos e direcionar as estratégias de manutenção.

Através da utilização de algoritmos de inteligência artificial, foi possível desenvolver um programa capaz de analisar dados de sensores e prever com exatidão falhas nos equipamentos. Os resultados obtidos demonstraram a eficácia deste programa na detecção de falhas de bateria nos seguidores fotovoltaicos, permitindo a substituição tempestiva e evitando períodos prolongados de inatividade. Essa intervenção imediata contribui para a maximização do tempo de operação dos sistemas fotovoltaicos, resultando em maior produção de energia e melhor desempenho global.

O modelo LSTM multivariável, que utilizou todas as variáveis como entradas, apresentou resultados promissores na previsão do estado de carga da bateria. A exatidão média para as baterias avariadas foi de 86,11%, enquanto para as baterias funcionais foi de 42,16%. Isso realça a capacidade do modelo em discernir entre falhas temporárias, causadas por anomalias ou condições climáticas desfavoráveis devido a baixa exposição solar, e falhas que exigem intervenção imediata devido à incapacidade de retorno ao funcionamento normal. Essa distinção possibilita uma gestão mais eficaz dos recursos de manutenção,



direcionando as intervenções apenas para os casos que demandam atenção imediata e otimizando a alocação de recursos.

Por sua vez, o modelo LSTM univariável, que considerou apenas uma variável como entrada, também obteve resultados satisfatórios na previsão do estado de carga da bateria.

A exatidão média para as baterias avariadas foi de 87,50%, enquanto para as baterias funcionais foi de 54,54%. Estes resultados demonstram a capacidade do modelo em detetar a necessidade de substituição da bateria a partir do primeiro alarme de falha, possibilitando uma intervenção rápida e eficaz.

A simulação económica realizada evidenciou as perdas energéticas e económicas significativas decorrentes de falhas nos seguidores fotovoltaicos. Esta análise reforça a importância de um programa de monitorização e manutenção dos sistemas, que permita a deteção antecipada de falhas e ações corretivas imediatas. Esta abordagem pode reduzir o tempo de inatividade dos equipamentos, minimizar os custos de reparação e maximizar a produção de energia.

Este estudo contribui para a compreensão do potencial da inteligência artificial na determinação do estado dos seguidores fotovoltaicos e destaca a relevância deste programa de manutenção preditiva para a maximização da produção de energia, a redução de custos operacionais e a gestão eficiente das centrais fotovoltaicas. À medida que se busca a conservação de energia e a ampliação do uso de fontes renováveis, a implementação de abordagens preditivas baseadas em inteligência artificial mostra-se crucial para o futuro da manutenção destes sistemas.

## 6.2 Trabalhos Futuros

Como perspetiva de desenvolvimento futuro, recomenda-se a contínua melhoria dos modelos de inteligência artificial utilizados neste estudo. A aplicação de técnicas avançadas de *machine learning*, como *deep learning* e algoritmos CNN e outros de RNN, permitirá a criação de modelos mais complexos e exatos. Esses modelos serão capazes de identificar padrões e antecipar a ocorrência de avarias nos seguidores fotovoltaicos com ainda maior exatidão. Além disso, a incorporação de um maior número de variáveis relevantes para a determinação do estado dos seguidores fotovoltaicos enriquecerá a análise e melhorará a capacidade preditiva dos modelos.

Para aprimorar a eficácia dos modelos de previsão, é recomendado integrar dados históricos de manutenção e reparação dos equipamentos. A recolha e análise de informações sobre falhas anteriores, procedimentos de manutenção realizados e reparações efetuadas podem fornecer insights valiosos para a identificação de padrões de falhas recorrentes e o desenvolvimento de estratégias de manutenção mais eficientes. Além disso, a integração destes dados com os modelos de inteligência artificial permitirá um ajuste contínuo dos algoritmos e uma maior capacidade de deteção precoce de falhas.

Como desenvolvimento futuro, sugere-se a criação de um programa capaz de prever antecipadamente as avarias das baterias dos seguidores fotovoltaicos, utilizando modelos

de inteligência artificial ainda mais avançados combinando a este modelo que consegue fazer a classificação objetiva entre baterias avariadas e não avariadas, e utilizar vários outros parâmetros que não foram possíveis utilizar pois não estavam disponíveis com a as tensões e cargas de descarga e recarga da bateria. Essa evolução permitirá uma gestão ainda mais eficiente dos sistemas fotovoltaicos, contribuindo para a sustentabilidade energética e o progresso contínuo das energias renováveis. A implementação deste programa possibilitará uma tomada de decisão mais proativa, permitindo intervenções preventivas e a minimização de tempos de paralisação dos equipamentos. Além disso, a monitorização constante e a previsão de avarias contribuirão para a otimização dos recursos de manutenção e a maximização da vida útil dos seguidores fotovoltaicos.

Por fim, é importante considerar a viabilidade económica e a escalabilidade dos modelos propostos. Uma análise mais aprofundada do custo-benefício dos sistemas de manutenção preditiva baseados em inteligência artificial, levando em conta os custos de implementação, integração de dados e recursos computacionais, é necessária para avaliar sua viabilidade em diferentes contextos e escalas de implementação. Além disso, considerações sobre a disponibilidade de recursos e a capacidade de dimensionamento dos modelos são essenciais para garantir que as soluções propostas possam ser implementadas de forma eficiente e sustentável.

Em conclusão, a continuidade da pesquisa e do desenvolvimento nos trabalhos futuros mencionados acima permitirá avanços significativos no campo da determinação do estado de seguidores fotovoltaicos com o uso de inteligência artificial. Essas melhorias contribuirão para aprimorar a eficiência e a fiabilidade dos sistemas fotovoltaicos, maximizando o desempenho energético, reduzindo custos operacionais e impulsionando ainda mais o progresso das energias renováveis

## BIBLIOGRAFIA

- Agrawal, A., Gans, J., & Goldfarb, A. (2016). The simple economics of machine intelligence. *Harvard Business Review*, 17(1), 2–5 (ver p. 11).
- Ali, A., & Abdelhadi, A. (2022). Condition-based monitoring and maintenance: state of the art review. *Applied Sciences*, 12(2), 688 (ver pp. 10, 11).
- Alpaydin, E. (2020). *Introduction to machine learning*. MIT press. (Ver p. 16).
- Bartoli, A., & Fusiello, A. (2021). *Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part VI* (Vol. 12540). Springer Nature. (Ver pp. 21, 22).
- Basri, E. I., Abdul Razak, I. H., Abdul Samat, H., & Kamaruddin, S. (2017). Preventive Maintenance (PM) planning: a review. *Journal of Quality in Maintenance Engineering*, 23. <https://doi.org/10.1108/JQME-04-2016-0014> (ver pp. viii, 6–8).
- Benedetti, M. L. D. (2018). Anomaly detection and predictive maintenance for photovoltaic systems. *Neurocomputing*, 310, 59–68. <https://doi.org/10.1016/j.neucom.2018.05.017> (ver p. 28).
- Betti, A., Trovato, M. L. L., Leonardi, F. S., Leotta, G., Ruffini, F., & Lanzetta, C. (2019). Predictive maintenance in photovoltaic plants with a big data approach. *arXiv preprint arXiv:1901.10855* (ver pp. viii, 27).
- Bosman, L. B., & Darling, S. B. (2018). Performance modeling and valuation of snow-covered PV systems: examination of a simplified approach to decrease forecasting error. *Environmental Science and Pollution Research*, 25, 15484–15491 (ver p. 26).
- Bosman, L. B., Leon-Salas, W. D., Hutzler, W., & Soto, E. A. (2020). PV system predictive maintenance: Challenges, current approaches, and opportunities. *Energies*, 13(6), 1398 (ver p. 26).
- Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). Predictive Maintenance in the 4th Industrial Revolution: Benefits, Business Opportunities, and Managerial Implications. *IEEE Engineering Management Review*, 48, 57–62. <https://doi.org/10.1109/EMR.2019.2958037> (ver p. 9).
- Cakir, M., Guvenc, M. A., & Mistikoglu, S. (2021). The experimental application of popular machine learning algorithms on predictive maintenance and the design of IIoT

- based condition monitoring system. *Computers Industrial Engineering*, 151, 106948. <https://doi.org/https://doi.org/10.1016/j.cie.2020.106948> (ver p. 9).
- Cardoso, D., & Ferreira, L. (2021). Application of Predictive Maintenance Concepts Using Artificial Intelligence Tools. *Applied Sciences*, 11(1). <https://doi.org/10.3390/app11010018> (ver p. 11).
- Chen, J. C., Chen, T.-L., Liu, W.-J., Cheng, C., & Li, M.-G. (2021). Combining empirical mode decomposition and deep recurrent neural networks for predictive maintenance of lithium-ion battery. *Advanced Engineering Informatics*, 50, 101405 (ver pp. viii, 28, 30).
- Chen, Z., Wu, L., Cheng, S., Lin, P., Wu, Y., & Lin, W. (2017). Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and IV characteristics. *Applied energy*, 204, 912–931 (ver pp. 26, 27).
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2015). Gated feedback recurrent neural networks. *International conference on machine learning*, 2067–2075 (ver p. 17).
- Çınar, Z. M., Abdussalam Nuhu, A., Zeeshan, Q., Korhan, O., Asmael, M., & Safaei, B. (2020). Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0. *Sustainability*, 12(19). <https://doi.org/10.3390/su12198211> (ver pp. 7, 25).
- Contents. (2002). Em R. K. Mobley (Ed.), *An Introduction to Predictive Maintenance (Second Edition)* (Second Edition, pp. v–xii). Butterworth-Heinemann. <https://doi.org/https://doi.org/10.1016/B978-075067531-4/50000-2> (ver pp. 6, 9).
- De Benedetti, M., Leonardi, F., Messina, F., Santoro, C., & Vasilakos, A. (2018). Anomaly detection and predictive maintenance for photovoltaic systems. *Neurocomputing*, 310, 59–68 (ver pp. 26, 28).
- Dixit, P., & Silakari, S. (2022). Analysis of state-of-art attack detection methods using recurrent neural network. *Proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences: PCCDS 2021*, 795–804 (ver p. 16).
- El Naqa, I., & Murphy, M. J. (2015). *What is machine learning?* Springer. (Ver p. 11).
- Esen, H., Inalli, M., Sengur, A., & Esen, M. (2008). Performance prediction of a ground-coupled heat pump system using artificial neural networks. *Expert Systems with Applications*, 35(4), 1940–1948 (ver p. 16).
- Gackowiec, P. (2019). General overview of maintenance strategies – concepts and approaches. *Multidisciplinary Aspects of Production Engineering*, 2, 126–139. <https://doi.org/10.2478/mape-2019-0013> (ver pp. 6, 7).
- Kaparthi, S., & Bumblauskas, D. (2020). Designing predictive maintenance systems using decision tree-based machine learning techniques. *International Journal of Quality & Reliability Management*, 37(4), 659–686 (ver pp. 13, 14).
- Krogh, A. (2008). What are artificial neural networks? *Nature biotechnology*, 26(2), 195–197 (ver p. 16).

- Livera, A., Theristis, M., Charalambous, A., Stein, J. S., & Georghiou, G. E. (2021). Decision support system for corrective maintenance in large-scale photovoltaic systems. *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)*, 0306–0311 (ver p. 10).
- Murat, Y. S., & Ceylan, H. (2006). Use of artificial neural networks for transport energy demand modeling. *Energy policy*, 34(17), 3165–3172 (ver p. 16).
- Nielsen, M. A. (2015). *Neural networks and deep learning* (Vol. 25). Determination press San Francisco, CA, USA. (Ver pp. viii, 17).
- O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458* (ver pp. viii, 16, 22).
- Osisanwo, F., Akinsola, J., Awodele, O., Hinmikaiye, J., Olakanmi, O., Akinjobi, J., et al. (2017). Supervised machine learning algorithms: classification and comparison. *International Journal of Computer Trends and Technology (IJCTT)*, 48(3), 128–138 (ver p. 15).
- Ouahdah, A., Zemmouchi-Ghomari, L., & Salhi, N. (2022). Selecting an appropriate supervised machine learning algorithm for predictive maintenance. *The International Journal of Advanced Manufacturing Technology*, 119(7-8), 4277–4301 (ver pp. viii, 13–15).
- Piramuthu, S., Raman, N., & Shaw, M. J. (1998). Decision support system for scheduling a flexible flow system: incorporation of feature construction. *Annals of Operations Research*, 78(0), 219–234 (ver p. 13).
- Rivas, A., Fraile, J. M., Chamoso, P., González-Briones, A., Sittón, I., & Corchado, J. M. (2020). A predictive maintenance model using recurrent neural networks. *14th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2019) Seville, Spain, May 13–15, 2019, Proceedings 14*, 261–270 (ver p. 18).
- Rokach, L., & Maimon, O. (2005). Top-down induction of decision trees classifiers-a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 35(4), 476–487 (ver pp. viii, 13).
- Schmidt, R. M. (2019). Recurrent neural networks (rnns): A gentle introduction and overview. *arXiv preprint arXiv:1912.05911* (ver p. 17).
- Schneider, H. (1996). Failure mode and effect analysis: FMEA from theory to execution. (Ver p. 10).
- Silvestre, S., da Silva, M. A., Chouder, A., Guasch, D., & Karatepe, E. (2014). New procedure for fault detection in grid connected PV systems based on the evaluation of current and voltage indicators. *Energy Conversion and Management*, 86, 241–249 (ver pp. 26, 27).
- Simões, J., F. Gomes, C., & Yasin, M. (2011). A literature review of maintenance performance measurement: A conceptual framework and directions for future research. *Journal of Quality in Maintenance Engineering*, 17, 116–137. <https://doi.org/10.1108/13552511111134565> (ver p. 7).

- Smagulova, K., & James, A. P. (2019). A survey on LSTM memristive neural network architectures and applications. *The European Physical Journal Special Topics*, 228(10), 2313–2324 (ver p. 18).
- Söderholm, P., Holmgren, M., & Klefsjö, B. (2007). A process view of maintenance and its stakeholders. *Journal of Quality in Maintenance Engineering*, 13, 19–32. <https://doi.org/10.1108/13552510710735096> (ver p. 8).
- Wang, K., & Wang, Y. (2018). How AI affects the future predictive maintenance: a primer of deep learning. *Advanced Manufacturing and Automation VII* 7, 1–9 (ver pp. 12, 16).
- Wang, Y., Deng, C., Wu, J., Wang, Y., & Xiong, Y. (2014). A corrective maintenance scheme for engineering equipment. *Engineering Failure Analysis*, 36, 269–283. <https://doi.org/10.1016/j.engfailanal.2013.10.006> (ver p. 10).
- Xie, J., Huang, J., Zeng, C., Jiang, S.-H., & Podlich, N. (2020). Systematic literature review on data-driven models for predictive maintenance of railway track: Implications in geotechnical engineering. *Geosciences*, 10(11), 425 (ver p. 22).
- Young, M. J. (2003). Human performance model validation: one size does not fit all. *Em Summer Computer Simulation Conference* (pp. 732–735). (Ver p. 8).
- Zbiciak, A., & Markiewicz, T. (2023). A new extraordinary means of appeal in the Polish criminal procedure: the basic principles of a fair trial and a complaint against a cassatory judgment. *Access to Justice in Eastern Europe*, 6(2), 1–18 (ver pp. viii, 19).
- Zhang, A., & Lipton, Z. C. (2023). *Dive into Deep Learning*. Cambridge University Press. (Ver pp. 19, 20).
- Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2021). Dive into deep learning. *arXiv preprint arXiv:2106.11342* (ver pp. 17, 18).
- Zhang, W., Yang, D., & Wang, H. (2019). Data-driven methods for predictive maintenance of industrial equipment: A survey. *IEEE systems journal*, 13(3), 2213–2227 (ver pp. 13, 23).
- Zonta, T., Da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the Industry 4.0: A systematic literature review. *Computers & Industrial Engineering*, 150, 106889 (ver p. 16).

