



Evento: XXXIII Seminário de Iniciação Científica ▾

APLICAÇÃO DE MODELOS MATEMÁTICOS NA PREDIÇÃO DO TEMPO DE VIDA DE BATERIAS QUE ALIMENTAM DISPOSITIVOS MÓVEIS¹

Leonardo Ojczenasz Schmidt², Airam Teresa Zago Romcy Sausen^{2,3}, Julia Dammann⁴, Jean Schäffel Moreira⁵

¹ Pesquisa desenvolvida na Unijuí; financiado pelo Programa Institucional de Bolsas de Iniciação Científica do Conselho Nacional de Desenvolvimento Científico e Tecnológico - PIBIC/CNPq

² Bolsista; estudante do curso Ciência da Computação; Bolsista do programa de fomento: Programa Institucional de Bolsas de Iniciação Científica do Conselho Nacional de Desenvolvimento Científico e Tecnológico - PIBIC/CNPq

³ Professor(a) Dra. orientador(a) do projeto Aplicação de Modelos Matemáticos na Predição do Tempo de Vida de Baterias que Alimentam Dispositivos Móveis.

⁴ Doutoranda em Modelagem Matemática e Computacional da UNIJUÍ.

⁵ Bolsista CNPq; estudante do curso Ciência da Computação da UNIJUÍ.

INTRODUÇÃO

A predição precisa do tempo de vida de baterias é um desafio central em dispositivos móveis, impactando diretamente sua autonomia e confiabilidade, e modelos matemáticos robustos são essenciais para essa tarefa. As Redes Neurais Artificiais (RNAs) surgem como uma abordagem proeminente devido à sua capacidade de aprender e modelar as complexas dinâmicas não lineares que caracterizam o processo de degradação eletroquímica. Este trabalho, desenvolvido no contexto de uma bolsa de iniciação científica (PIBIC/CNPq) e em apoio a uma Tese de Doutorado, concentra-se no aprimoramento de modelos preditivos para baterias de Lítio-Íon Polímero (Li-Po), utilizando um conjunto de dados experimentais obtidos em uma plataforma de testes do Grupo de Automação Industrial e Controle (GAIC).

A pesquisa seguiu uma metodologia bem definida, partindo da premissa de que o desempenho de um modelo de aprendizado de máquina é função tanto de sua arquitetura quanto da qualidade dos dados de treinamento. O estudo se iniciou com a otimização da arquitetura da rede, um passo fundamental para garantir que o modelo possuísse a complexidade necessária para capturar os padrões nos dados, contudo, ao se constatar que apenas alterações estruturais atingiam um limite de performance, a pesquisa avançou para uma segunda frente: a manipulação dos dados de treinamento.



Esta etapa visou suavizar um dos principais gargalos em projetos de aprendizado de máquina, a escassez de dados, através da aplicação de técnicas de ampliação artificial e para tal, foi explorada a técnica SMOTE (Synthetic Minority Oversampling Technique) (CHAWLA et al., 2002), e sua adaptação para problemas de regressão, o SMOTER (Synthetic Minority Oversampling Technique for Regression) (TORGO et al., 2013). Desta forma, a pesquisa se aprofundou na investigação, desenvolvimento e aplicação de uma técnica para gerar mais dados de treinamento, com o objetivo de fornecer mais exemplos para a rede, aprimorar sua capacidade de generalização e, conseqüentemente, reduzir o erro de predição.

METODOLOGIA

A abordagem metodológica foi estruturada em três pilares: fundamentação teórica, otimização de arquitetura e implementação da ampliação de dados. Inicialmente, uma pesquisa bibliográfica aprofundada foi realizada, com destaque para a Dissertação de Mestrado de Dammann (2022), que forneceu o contexto de pesquisa, e os artigos de Chawla et al. (2002) e Torgo et al. (2013), que foram fundamentais para a compreensão teórica das técnicas de sobreamostragem.

A primeira fase experimental, conduzida no software MATLAB, foi dedicada à otimização da arquitetura da rede neural e, com o objetivo de encontrar a topologia mais adequada para o problema, foram realizados testes sistemáticos com 9 estruturas de rede distintas. Essas estruturas variavam em profundidade (uma ou duas camadas ocultas) e em largura (número de neurônios por camada), permitindo uma análise detalhada de como a complexidade do modelo influencia sua performance.

A segunda fase experimental concentrou-se na ampliação do conjunto de dados, onde na análise prévia já foi possível constatar que não poderia ser utilizada a técnica SMOTE, pois em sua forma original, ela é incompatível com problemas de regressão, pois opera sobre problemas de classificação.

Diante disso, foi pesquisado e selecionado a técnica SMOTER, que é uma adaptação do SMOTER para problemas de regressão. Foi desenvolvido um script em Python baseado em um ecossistema de bibliotecas robustas: Pandas e Openpyxl foram utilizados para a leitura

e manipulação dos dados de entrada e saída; NumPy foi empregado para as operações vetoriais e matriciais de alta performance; a biblioteca Scikit-Learn forneceu a implementação do algoritmo NearestNeighbors, que é o cerne do SMOTER para encontrar os vizinhos mais próximos de cada amostra; e, por fim, a biblioteca SciPy foi utilizada para salvar os dados sintéticos gerados no formato .mat, garantindo a perfeita compatibilidade com o ambiente MATLAB para a etapa de treinamento. A lógica do SMOTER consiste em gerar novos pontos através de uma interpolação linear entre uma amostra existente e um de seus vizinhos mais próximos. A nova amostra é gerada pela equação:

$$\vec{x}_{\text{novo}} = \vec{x}_i + \delta \cdot (\vec{x}_j - \vec{x}_i)$$

onde: x_i é a amostra original, x_j é o vizinho escolhido aleatoriamente e δ é um número aleatório no intervalo $[0, 1]$. Este mesmo processo de interpolação é aplicado de forma análoga à variável de saída (a capacidade da bateria), uma etapa crucial para garantir que as amostras sintéticas preservem as características estatísticas e a correlação entre entrada e saída presentes na vizinhança local dos dados originais.

RESULTADOS E DISCUSSÃO

A primeira etapa da pesquisa, focada na identificação da estrutura de rede neural mais otimizada, produziu um resultado claro sendo testadas 9 combinações de arquiteturas, variando o número de camadas e a distribuição de neurônios.

Avaliando o desempenho das redes pelo coeficiente de determinação no conjunto completo (treino, validação e teste), observou-se que, apesar dos resultados competitivos, a Rede G (duas camadas ocultas de 10 neurônios) teve a melhor acurácia, com 94,58%. Isso sugere que a maior profundidade ajudou a capturar melhor os padrões de degradação, levando à escolha dessa arquitetura como base para as próximas etapas.

Com a arquitetura otimizada, a pesquisa avançou para a aplicação da técnica SMOTER, onde foi implementado e executado um script em Python com sucesso, gerando os arquivos .mat com os dados ampliados. Para validar a abordagem, a rede neural foi treinada



em duas condições distintas: uma utilizando apenas os dados originais e outra utilizando o conjunto de dados aumentado, sendo a análise dos resultados dividida em duas frentes, que revelaram um comportamento ambíguo e de grande interesse científico.

A primeira análise comparou a capacidade da rede de aprender o comportamento de um modelo matemático teórico de referência (modelo de Kim), que descreve um processo de degradação idealizado e nesta comparação, a rede treinada com os dados sintéticos do SMOTER obteve um desempenho notavelmente superior, com um erro médio de apenas 3,22%, em contraste com o erro de 7,71% da rede treinada com os dados originais. Este resultado, que representa uma redução relativa de aproximadamente 58% no erro, indica que o SMOTER foi altamente eficaz em fazer a rede se ajustar a uma função matemática "limpa". A hipótese é que a interpolação inerente ao SMOTER cria uma distribuição de dados mais densa e suave, o que facilita o aprendizado de uma função teórica, que por definição não possui o ruído e as anomalias de dados experimentais.

A segunda análise, por sua vez, avaliou a capacidade da rede de prever o comportamento real da bateria, comparando suas previsões com os dados experimentais (Tvem) e neste cenário, que representa a aplicação prática do modelo, o resultado se inverteu drasticamente. A rede treinada com os dados originais obteve um desempenho significativamente melhor, com um erro médio de 4,68%, em contrapartida, a rede treinada com os dados ampliados pelo SMOTER apresentou um erro médio de 9,87%, mais que o dobro da versão original. Este resultado contra intuitivo é a principal contribuição do trabalho, pois evidencia que a técnica de ampliação de dados, embora benéfica em um cenário idealizado, foi prejudicial no cenário real.

CONSIDERAÇÕES FINAIS

Durante o projeto, um dos primeiros desafios técnicos foi a constatação de que a técnica SMOTE, amplamente utilizada em classificação, não era diretamente aplicável ao problema de regressão, o que exigiu uma pesquisa por alternativas e resultou na adoção e implementação do SMOTER. A integração entre as ferramentas Python e MATLAB, embora



bem-sucedida, também representou uma etapa que demandou atenção aos detalhes de compatibilidade de dados.

Contudo, a principal dificuldade é o achado mais significativo do trabalho surgiram da análise dos resultados, pois foi inesperado observar que a rede treinada com os dados gerados pelo SMOTER obteve um desempenho superior na predição do modelo teórico, mas, ao mesmo tempo, um desempenho pior na predição dos dados experimentais que são a base de todo o estudo. Esta dualidade de resultados leva a crer que a técnica SMOTER, ao gerar novos pontos por interpolação, pode ter suavizado excessivamente a distribuição dos dados, eliminando o ruído e as pequenas variações que, no mundo real, são informações cruciais sobre o processo de degradação da bateria. Este resultado sugere que a técnica, no formato aplicado, pode não ser a mais adequada para este problema específico, ou que talvez necessite de uma calibração mais detalhada. Este estudo, portanto, não apenas cumpriu o objetivo de otimizar a rede, mas também revelou uma importante limitação da técnica de ampliação de dados no contexto estudado, reforçando uma máxima da ciência de dados: a necessidade de validação rigorosa contra dados do mundo real, pois o sucesso em um ambiente controlado não garante a eficácia em uma aplicação prática.

Palavras-chave: Redes Neurais Artificiais. Modelagem Matemática. Baterias. SMOTER. Aprendizado de Máquina.

REFERÊNCIAS BIBLIOGRÁFICAS

CHAWLA, N. V. et al. Smote: synthetic minority over-sampling technique. **Journal of artificial intelligence research**, v. 16, p. 321-357, 2002.

DAMMANN, Julia. **Modelagem Matemática do Tempo de Vida de Baterias Através da Proposição de um Modelo Híbrido Personalizado que Considera o Número de Ciclos de Descarga**. 2022. Dissertação (Mestrado) - Universidade Regional do Noroeste do Estado do Rio Grande do Sul - UNIJUÍ, Ijuí. 2022.

TORGO, L. et al. Smote for regression. In: SPRINGER. **European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases**. 2013. p. 348-363.