

Modelo de aprendizado de máquina não supervisionado para detecção de falha em processos e equipamentos

Victor Zoratti Ferreira¹, Felipe Augusto Rodrigues Perez², Eduardo Soares Zanutti³, Junior Fernandes Marques⁴, Érico dos Reis⁵, Eduardo Brito⁶, Edfram Rodrigues Pereira⁷, Francisco Louzada Neto⁸

ICMC-USP

1 Introdução

A detecção de anomalias é um tema relevante tanto em termos acadêmicos quanto em aplicações práticas, pois sua identificação precoce pode prevenir falhas críticas e aumentar a eficiência operacional nas indústrias. No entanto, a maioria dos dados disponíveis refere-se ao funcionamento normal das máquinas, limitando o uso de técnicas preditivas supervisionadas. Nesse contexto, métodos de aprendizagem profunda, como os autoencoders (AE), demonstram-se promissores por permitirem uma abordagem não supervisionada em cenários com poucos dados de falha, possibilitando uma detecção mais robusta [1].

Embora muitos estudos utilizem dados simulados e ambientes controlados para validar suas hipóteses, a aplicação em dados reais torna-se fundamental para uma avaliação precisa e contundente, pois os ambientes industriais podem apresentar desafios inerentes de casos práticos como a escassez de dados rotulados, outliers e ruídos, que precisam ser abordados e tratados de forma correta para garantir a eficácia das técnicas de detecção de anomalias [2].

O presente artigo buscou avaliar a performance deste tipo de modelagem na detecção de anomalias, com o objetivo de contribuir para a manutenção preditiva e a eficiência operacional em sistemas de bombeamento industrial, em especial em motobombas da Petrobras Transporte S/A (TRANSPETRO).

¹victor.zoratti.ferreira@usp.br

²felipe.perez@usp.br

³eduardozanutti@usp.br

⁴junior.marques@usp.br

⁵ericodosreis@usp.br

⁶ebritojr@usp.br

⁷epereira@usp.br

⁸louzada@icmc.usp.br

2 Metodologia

2.1 Autoencoder

Um autoencoder (AE) é um modelo de rede neural artificial que aprende padrões de forma eficiente por meio de camadas conhecidas como encoder, que comprimem os dados de entrada, reduzindo sua dimensionalidade. Posteriormente, um segundo conjunto de camadas (decoder) tenta reconstruir o sinal inicial comprimido. A diferença entre eles, original e reconstruído, é conhecida como erro de reconstrução, é uma medida de dissimilaridade que pode indicar possíveis anomalias na série histórica. Sendo assim, trata-se de um modelo de aprendizado não supervisionado, pois ele aprende padrões sem necessidade de rótulos de anomalia [3].

2.2 Aplicação ao Tennessee Eastman Process (TEP)

Este banco de dados é amplamente utilizado na literatura para monitoramento de processos químicos e origina-se da simulação desenvolvida na Companhia Eastman, com sede em Kingsport, Tennessee para experimentos de classificação de falhas em equipamentos industriais, uma vez que anomalias sintéticas são incluídas em algumas máquinas de maneira aleatória [4].

2.3 Aplicação ao caso real da TRANSPETRO

Diversas variáveis de processo, dos equipamentos operados pela TRANSPETRO, são monitoradas e armazenadas como séries temporais multivariadas. Dentre elas, corrente, vazão, pressão de sucção, descarga, corrente elétrica, temperatura e vibração do motor e da bomba em diferentes pontos de medição foram utilizadas para criar o AE de uma motobomba que apresentou falhas no passado.

2.4 Fluxo do trabalho

Este trabalho foi desenvolvido utilizando a linguagem de programação PYTHON. Primeiramente, foi elaborado um AE para os dados do TEP, onde o objetivo foi provar que o modelo conseguiria detectar as anomalias sintéticas geradas pelo conjunto de dados simulados. Posteriormente, uma análise de correlações nos dados da TRANSPETRO ajudou na seleção das variáveis significantes. Por fim, a Figura (1) explica o passo a passo da implementação do AE.

As Equações (1) e (2) foram utilizadas para calcular o Grau de Anomalia (GA) e Mínimo Móvel Grau de Anomalia (MMGA) respectivamente, métricas úteis para detectar em que momento as falhas ocorrem. Uma janela de 10 períodos foi utilizada, e o limiar definido como 1,5 vezes o valor máximo do MMGA. Assim, se após 10 ocorrências acima da referência, o modelo classificará como anomalia, pois entende-se que a inconsistência é relevante na série histórica e portanto passível de parada técnica e/ou manutenção preditiva.

$$GA(X^{(i)}) = \left\| \tilde{X}^{(i)} - X^{(i)} \right\|^2, \quad (1)$$

$$MMGA(X^{(i)}) = \min \left(GA(X^{(i)}), GA(X^{(i-1)}), \dots, GA(X^{(i-9)}) \right), \quad (2)$$

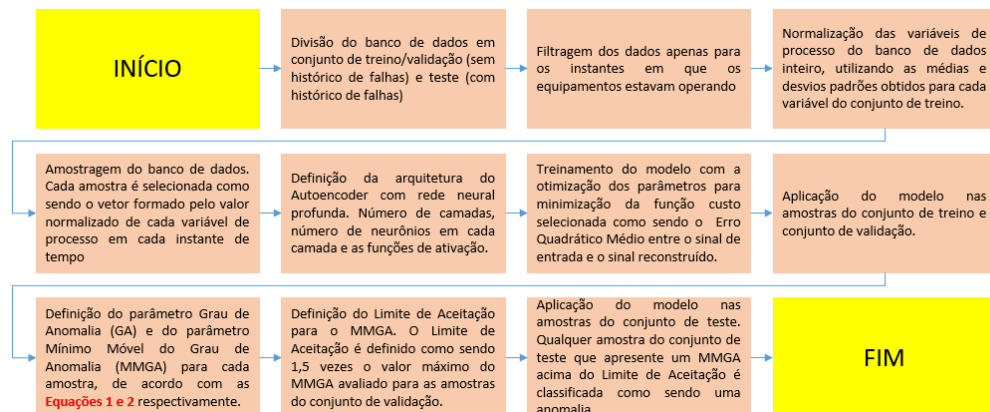


Figura 1: Fluxograma Metodologia.

3 Resultados

A seguir, são apresentados os resultados obtidos dos modelos de AE para a detecção de anomalias, tanto no banco de dados simulado da TEP, quanto na base do equipamento real operado pela TRANSPETRO.

3.1 TEP

Podemos observar na figura (2) que o erro de reconstrução ultrapassa o limite de aceitação do modelo a partir da inserção da falha sintética e, portanto, detectando-a como anômala.

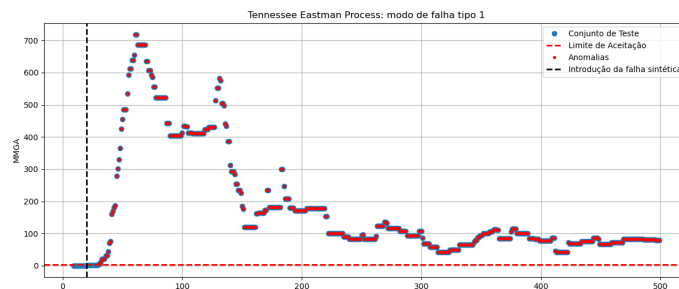


Figura 2: Falha detectada do tipo 1

A partir dessa experimentação, o mesmo método foi utilizado para os equipamentos reais da TRANSPETRO.

3.2 TRANSPETRO

De maneira semelhante, o AE detectou a presença de inconsistências na reconstrução dos dados, o limite de aceitação demonstrou que a anomalia vinha ocorrendo um pouco antes da falha

identificada em campo, mas somente após um determinado período ela ficou ainda mais evidente e captada pela equipe operacional como nota-se na Figura (3).

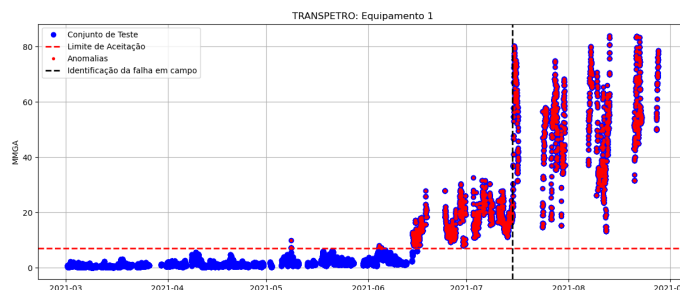


Figura 3: Falha detectada equipamento 1

4 Resultados

Os resultados preliminares deste trabalho evidenciam o potencial de modelos de aprendizado de máquina não supervisionado, com ênfase em redes neurais, para a detecção de anomalias em processos e equipamentos industriais. Ele foi desenvolvido na disciplina MAI5003 - Probabilidade e Estatística, utilizando a abordagem de Aprendizado Baseado em Problemas (PBL).

Nesta pesquisa, foi elaborado um modelo de AE com rede neural profunda, utilizando a linguagem de programação PYTHON. Para estudos futuros, outros tipos de modelos e técnicas podem ser explorados, como autoencoders convolucionais e recorrentes, que podem capturar diferentes padrões de comportamento em séries temporais.

Além dos autoencoders, outros modelos de detecção de anomalias também podem ser testados e avaliados para a análise de séries multivariadas. Por fim, esse estudo contribui para a literatura ao expandir o uso de autoencoders na detecção de anomalias em motobombas de óleo e petróleo com dados reais propondo uma metodologia eficaz para a solução deste tipo de problema.

Referências

- [1] G. S. Chadha, A. Rabbani and A. Schwung. Comparison of Semi-supervised Deep Neural Networks for Anomaly Detection in Industrial Processes, *Int. Con. On Ind. Inf.*, 2019. DOI: 10.1109/INDIN41052.2019.8972172.
- [2] A. C. O. e Souza, M. B. de Souza Jr. and F. V. da Silva. Development of a CNN-based fault detection system for a real water injection centrifugal pump, *Exp. Syst with App.*, 2024. DOI: 10.1016/j.eswa.2023.122947.
- [3] D. E. Sapina et al. Comparison of autoencoder architectures for fault detection in industrial processes, *Dig. Chem. Enge.*, 2024. DOI: 10.1016/j.dche.2024.100162.
- [4] J.J. Downs and E.F. Vogel. A plant-wide industrial process control problem, *Comp. Chem. Enge.*, 1993. DOI: 10.1016/0098-1354(93)80018-I.