

APLICAÇÃO DE APRENDIZADO DE MÁQUINA PARA DETECÇÃO DE FALHAS EM SISTEMAS DE ELEVAÇÃO DE PETRÓLEO

Kaio H. F. ROCHA ; Eduardo N. de OLIVEIRA ; Rejane B. SANTOS

RESUMO

Este trabalho apresenta o desenvolvimento de um modelo de aprendizado de máquina para detecção automática de falhas em sistemas de elevação de petróleo. Utilizando o banco de dados 3W, foi aplicada uma rede neural treinada com dados operacionais tratados e normalizados. A arquitetura implementada demonstrou desempenho elevado, alcançando 96,43% de acurácia na fase de testes. A matriz de confusão evidenciou baixa taxa de erros, indicando alta confiabilidade na classificação dos eventos. Os resultados reforçam o potencial da inteligência artificial como ferramenta de apoio ao monitoramento e à manutenção preditiva em ambientes industriais complexos.

Palavras-chave:

Poços produtores de petróleo; detecção de falhas; Aprendizado de Máquina.

1. INTRODUÇÃO

A indústria de petróleo e gás enfrenta desafios constantes relacionados à operação de seus sistemas, frequentemente submetidos a condições extremas de pressão, temperatura e vazão. A ocorrência de eventos indesejados, como falhas operacionais e anomalias em equipamentos, pode resultar em prejuízos econômicos significativos e riscos à segurança das instalações (Sobrinho et al., 2020).

A integração de sensores de pressão e temperatura permite a análise contínua do desempenho do poço, possibilitando a identificação de anomalias que possam comprometer sua integridade (Sobrinho et al., 2020). A detecção precoce e a correta classificação desses eventos são fundamentais para prevenir falhas e orientar a tomada de decisões (Machado et al., 2022).

Nesse contexto, o uso de técnicas baseadas em aprendizado de máquina tem se destacado como uma alternativa promissora para o monitoramento inteligente e a detecção precoce de falhas em sistemas industriais complexos (Machado et al., 2022). Tais métodos possibilitam o aprimoramento da eficiência operacional, além de reduzir interrupções não planejadas e custos associados à manutenção corretiva (Júnior et al., 2020).

Considerando a relevância do tema, o presente trabalho tem como objetivo aplicar técnicas de aprendizado de máquina na detecção de falhas em sistemas de extração de petróleo, mais especificamente os eventos classificados como *Severe Slugging* e *Flow Instability*. Para isso, será utilizado o banco de dados 3W Dataset, desenvolvido por Vargas et al. (2019), que reúne observações reais de diferentes poços de petróleo.

2. MATERIAL E MÉTODOS

O material em foco nesta pesquisa é o banco de dados criado por Vargas *et al.* (2019), chamado *3W dataset*, que contém registros sobre variáveis de operação de plataformas de petróleo da Petrobras. Neste compilado, encontram-se dados sobre o estado de cada uma dessas variáveis durante o funcionamento cotidiano e durante a ocorrência de falhas específicas, denominadas pelo autor como eventos indesejados.

Neste estudo, foram analisadas duas falhas principais: *Severe Slugging*, caracterizado por oscilações abruptas de pressão e vazão, comprometendo a estabilidade operacional, e *Flow Instability*, definido por instabilidades menos intensas, porém persistentes no escoamento, podendo originar o primeiro evento citado. Ambos são de grande relevância prática, uma vez que impactam diretamente a segurança, a confiabilidade e a continuidade das operações em sistemas de elevação de petróleo.

Com base no comportamento registrado pelos sensores, foi desenvolvido um algoritmo de aprendizado de máquina capaz de identificar a ocorrência de falhas em tempo real, ao associar a divergência das variáveis em relação ao processo em estado regular. Para isso, os dados foram fornecidos a um algoritmo com a classificação indicativa da presença ou não de desvios operacionais.

O processo metodológico teve início com uma etapa de análise e preparação dos dados brutos. Cada variável foi avaliada individualmente, sendo identificados problemas comuns em dados industriais, como valores nulos, leituras congeladas geralmente associados a sensores defeituosos. Um tratamento específico foi aplicado à variável QGL, cujos valores nulos foram descartados, enquanto os valores iguais a zero foram mantidos como representações válidas de operação. Ademais, séries com 15 ou mais repetições consecutivas do mesmo valor foram consideradas como falhas de sensor e removidas. Após essa etapa, colunas com mais de 30% de registros inconsistentes foram eliminadas por completo, enquanto os demais dados inválidos foram descartados seletivamente.

Com o conjunto de dados limpo, procedeu-se à normalização das variáveis operacionais, utilizando a técnica *MinMaxScaler*. A padronização em uma escala de 0 a 1 permitiu que o modelo interpretasse adequadamente a relevância relativa de cada parâmetro durante o aprendizado, evitando o viés de magnitude entre variáveis distintas.

Na fase de modelagem, foi adotada uma arquitetura de rede neural artificial (*ANN – Artificial Neural Network*), com três camadas ocultas compostas por 16 neurônios cada. A função de ativação utilizada nas camadas ocultas foi a *ReLU*, por sua eficiência em acelerar o aprendizado e mitigar o problema do desvanecimento do gradiente. Para evitar *overfitting*, foram incorporadas camadas de *Dropout* com taxa de 30%, enquanto a camada final adotou ativação sigmoide, adequada à tarefa de classificação binária, isto é, determinar se há ou não ocorrência de falha.

O treinamento foi conduzido com o otimizador Adam e a função de perda *binary crossentropy*, sendo monitorado com a técnica de *Early Stopping*, a fim de interromper o treinamento assim que a performance de validação deixasse de melhorar, garantindo a retenção dos melhores pesos obtidos. O conjunto de dados foi dividido em 70% para treinamento e 30% para teste, permitindo uma avaliação imparcial do desempenho do modelo em dados não vistos.

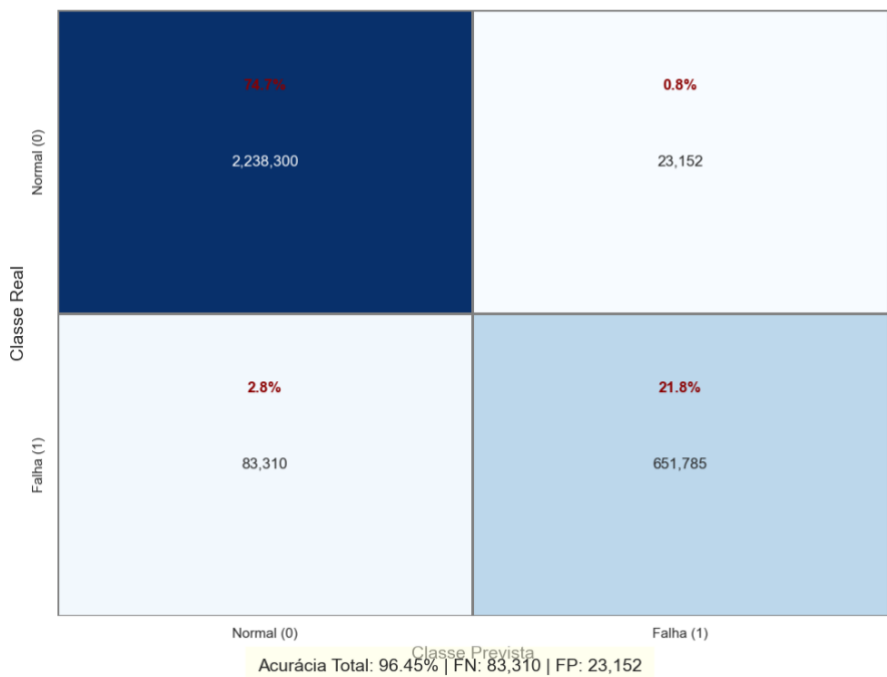
A avaliação do modelo foi realizada por meio de métricas clássicas de classificação binária, como acurácia, precisão, *F1-score* e matriz de confusão, com o objetivo de mensurar a capacidade do algoritmo em identificar corretamente as falhas e minimizar erros de classificação.

3. RESULTADOS E DISCUSSÃO

O modelo desenvolvido demonstrou alto desempenho na detecção de falhas em sistemas de elevação de petróleo, alcançando uma acurácia de 96,43%, indicando boa capacidade de identificar corretamente tanto eventos de falha quanto condições normais de operação.

A Figura 1 apresenta a matriz de confusão gerada a partir das previsões do modelo sobre o conjunto de teste. Nota-se que a maioria das amostras foi classificada corretamente, com baixa incidência de falsos positivos e negativos. Esse resultado evidencia a capacidade da rede neural em reconhecer padrões operacionais distintos com elevada confiabilidade.

Figura 01: Matriz de Confusão – Detecção de Falhas



Fonte: Elaborado pelo autor

4. CONCLUSÃO

Conclui-se, portanto, que o sistema proposto é eficaz na detecção automatizada de falhas em ambientes industriais complexos. Sua aplicação pode contribuir significativamente para o monitoramento inteligente de poços de petróleo, auxiliando na tomada de decisão e na redução de custos com manutenção corretiva e paradas não programadas.

REFERÊNCIAS

Júnior, W. F.; Vargas, R. E. V; Komati, K. S; Gazolli, K. A. de S. Detecção de anomalias em poços produtores de petróleo usando aprendizado de máquina. Sociedade brasileira de Automática, V. 02, N 01, 2020.

Machado, A. P. F.; Vargas, R. E. V; Ciarelli, P. M.; Munaro, C. J. Improving performance of one-class classifiers applied to anomaly detection in oil wells. Journal of Petroleum Science and Engineering, 218, 2022.

Sobrinho, E. S.P.; Oliveira, F. L.; Anjos, J. L.R.; Gonçalves, C.; Ferreira, M. V. D.; Lopes, L. G. O.; Lira, W. W. M.; Araújo, J. P. N.; Silva, T. B.; Gouveia, L. P. Uma ferramenta para detectar anomalias de produção utilizando aprendizagem profunda e árvore de decisão. Rio Oil & Gas Expo and Conference, 2020.

Vargas, R. Base de dados e Benchmarks para prognóstico de anomalias em sistemas de elevação de petróleo. Tese (Pós-graduação em Engenharia Elétrica) — Centro Tecnológico da Universidade Federal do Espírito Santo. UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO. Vitória, p. 75. 2019.

Vargas, R. E. V; Munuro, C. J.; Ciarelli, P.M.; Medeiros, A. G. M.; Amaral, B. G. do; Barrionuevo, D. C.; Araújo, J. C. D. de; Ribeiro, J. L.; Magalhães, L. P. A realistic and public dataset with rare undesirable real events in oil wells. Journal of Petroleum Science and Engineering, v. 181, 2019.