

Detecção de anomalia aplicado ao monitoramento da integridade estrutural em um sistema de tubulações.

Anomaly detection applied to structural health monitoring in a piping system.

Alexsander Lindolfo de Lima¹, Rodrigo Miguel Dias da Silva², Camila Pereira da Silva³, Renan Garcia Rosa⁴, Stanley Washington Ferreira de Rezende⁵, José dos Reis Vieira de Moura Júnior⁶

RESUMO

Neste trabalho, dois algoritmos do tipo semi-supervisionado foram comparados em relação a detecção de anomalias no Monitoramento da Integridade Estrutural baseado em Impedância (ISHM - *Impedance-based Structural Health Monitoring*) eletromecânica em um sistema de tubulações. O experimento consistiu em construir um sistema de tubulações de baixo custo com uma pastilha de Titanato Zirconato de Chumbo (PZT - *Pb-lead Zirconate Titanate*) acoplada na superfície principal com o auxílio de uma braçadeira, modelada e impressa em uma impressora 3D, capaz de envolver o tubo para melhorar a interação entre a pastilha PZT e a estrutura. Para simular situações de falhas, três tipos foram considerados, entupimento utilizando uma massa epóxi, dano provocado por riscados na parte interna do tubo e encrustamento de menor magnitude, provocado por um filtro dimensionado na parte interna do tubo. As técnicas escolhidas para realizar a classificação do estado natural e dano foram, Fator de Anomalia Local (LOF - *Local Outlier Factor*) e Máquina de Vetores de Suporte (SVM - *Support Vector Machine*), cada uma configurada de maneira semi-supervisionada. Utilizando as métricas de avaliação, Matriz de confusão e curva ROC, foi possível constatar que os dois métodos conseguiram identificar corretamente todas as classes do banco de dados.

Palavras-chave: Aprendizado semi-supervisionado. Detecção de anomalias. Monitoramento da Integridade Estrutural. Métricas de avaliação.

ABSTRACT

In this work, two semi-supervised algorithms were compared in relation to anomaly detection in electromechanical Impedance-based Structural Health Monitoring (ISHM) in a piping system. The experiment consisted of building a low-cost piping system with a Pb-lead Zirconate Titanate (PZT) patch attached to the main surface using a clamp, modeled and printed on a 3D printer, capable of wrapping around the pipe to improve the interaction between the PZT pellet and the structure. To simulate failure situations, three types were considered: clogging using an epoxy putty, damage caused by scratches on the inside of the pipe and minor fouling caused by a filter dimensioned on the inside of the pipe. The techniques chosen to classify the natural state and damage were Local Outlier Factor (LOF) and Support Vector Machine (SVM), each configured in a semi-supervised way. Using the evaluation metrics, confusion matrix and ROC curve, it was possible to see that the two methods were able to correctly identify all the classes in the dataset.

Keywords: Semi-supervised learning; Anomaly detection; Structural Health Monitoring; Evaluation Metrics.

¹ Bacharel em Matemática Industrial e Mestrando em Modelagem e Otimização, Universidade Federal de Catalão, ORCID: <https://orcid.org/0000-0002-9865-2781>.

E-mail:

alindolfo@discente.ufcat.edu.br

² Bacharel em Engenharia de Minas e Mestrando em Modelagem e Otimização, Universidade Federal de Catalão, ORCID: <https://orcid.org/0009-0009-1070-4662>.

³ Mestre em Modelagem e Otimização, Universidade Federal de Catalão, ORCID: <https://orcid.org/0000-0002-5626-5222>.

⁴ Mestre em Modelagem e Otimização, Universidade Federal de Catalão e Doutorando em Ciência da Computação, Universidade Federal de Uberlândia, ORCID: <https://orcid.org/0000-0002-4317-8680>.

⁵ Mestre em Modelagem e Otimização, Universidade Federal de Catalão e Doutorando em Engenharia Mecânica, Universidade Federal de Uberlândia, ORCID: <https://orcid.org/0000-0003-3163-9862>.

⁶ Doutor em Engenharia Mecânica, Universidade Federal de Catalão, ORCID: <https://orcid.org/0000-0002-8701-8355>.

1. INTRODUÇÃO

Nas últimas décadas, devido ao amplo avanço tecnológico, houve uma grande necessidade de métodos de manutenção mais sofisticados com o intuito de maximizar vida útil de equipamentos e minimizar perdas financeiras. Um destes métodos é o Monitoramento da Integridade Estrutural baseado em Impedância (ISHM - *Impedance-based Structural Health Monitoring*) que consiste em uma técnica não destrutiva da manutenção preditiva. Este procedimento pode ser implementado acoplado uma pastilha de Titanato Zirconato de Chumbo (PZT - *Pb-lead Zirconate Titanate*) na estrutura que, devido suas propriedades piezoelétricas, pode ser utilizado como sensor ou atuador, onde as assinaturas de impedância, adquiridas a partir da excitação em alta frequência da estrutura, são coletadas e em seguida comparadas com a finalidade de determinar sua natureza (GIURGIUTIU, ZAGRAI, 2005; LIANG, SUN, ROGERS 1994).

Por outro lado, técnicas de detecção de anomalias podem ser entendidas como identificação de padrões em dados que diferem da sua natureza original. Portanto, devido a diversidade de cenários em que se é aplicado estas técnicas, como por exemplo, detecção de fraude em sistemas financeiros, diagnósticos médicos e detecção de falhas em equipamentos industriais, diversos métodos foram desenvolvidos (BOUKERCHE, ZHENG, ALFANDI, 2020; CHANDOLA, BANERJEE, KUMAR, 2009). Dentre esses processos, diversos métodos baseados em estatística, densidade, distância, aprendizado de máquina e agrupamento foram projetados para atender diversas demandas em relação a reconhecimento de anomalias (WANG, BAH, HAMMAD, 2019).

Wang, Yang e Ahn propuseram um sistema de detecção de anomalias de energia com eficiência de amostragem (SEPAD - *Sample Efficient Power Anomaly Detection*) em Mkalama, zona rural localizada na Tanzânia. Neste sistema existem dois classificadores, o primeiro é um Classificador de Correspondência de Padrões de Aparelhos (APMC - *Appliance Pattern Matching Classifier*) e o segundo é um Classificador de Hábitos de Consumo de Energia (EHC - *Energy Consumption Habit Classifier*). O APMC consiste em um método de aprendizado semi-supervisionado utilizando máquinas de vetores de suporte e o EHC é baseado no modelo de Markov e tem como objetivo melhorar a racionalidade do SEPAD. Para demonstrar o desempenho do sistema, foi feita uma comparação entre o SEPAD utilizando função Gaussiana e Linear com outros três métodos, Máquina de Vetor de Suporte baseado em Detecção de Anomalias (SVM-AD - *Support Vector Machine based Anomaly Detector*), Média Móvel Integrada Auto-Regressiva junto com Redes Neurais

Artificiais baseadas em Detecção de Anomalias (ARIMA-ANN-AD - *Auto-Regressive Integrated Moving Average and Artificial Neural Networks based on Anomaly Detection*) e k-Vizinhos mais Próximos baseado em Detecção de Anomalias (kNN-AD - *k-Nearest Neighbors based Anomaly Detector*). Como resultado, o sistema SEPAD, utilizando a função Gaussiana, apresentou os melhores resultados em relação às métricas de avaliação, Acurácia e F_1 , ou seja, demonstrando uma elevada precisão em relação a detecção de anomalias (WANG, YANG, AHN, 2019).

Hussain, Du e Ren utilizaram um detector de anomalias baseado em estatística utilizando como premissa o aprendizado semi-supervisionado para analisar grandes volumes de dados gerados da rede 4G LTE-A, com o intuito de identificar uma região com baixo nível de atividade que pode representar uma célula (estação que fornece serviços de comunicação sem fio) adormecida, como também uma região de tráfego muito alto que requer uma alocação de recursos adicionais para evitar falhas. Por fim, foram utilizadas as métricas de avaliação, *Accuracy*, *Error Rate*, *F₁ Score*, *False Positive Rate (FPR)*, *Precision* e *Recall* para demonstrar a eficiência do método em relação a detecção de anomalias, ou seja, na redução dos custos da gestão e manutenção das redes (HUSSAIN, DU, REN, 2018).

Zemmal et al. utilizam o algoritmo Máquina de Vetores de Suporte (SVM - *Support Vector Machine*) semi-supervisionado para melhorar a precisão do diagnóstico assistido por computador (CAD - *Computer-Aided Diagnosis*) para o câncer de mama. Depois de feito à base de treinamento, extração das características e seleção dos recursos, os resultados demonstraram que esta técnica tem uma alta taxa de reconhecimento em relação a classificação de imagens advindas do CAD (ZEMMAL et al., 2016).

Ma, Ngan e Liu utilizaram o algoritmo Fator de Anomalia Local (LOF - *Local Outlier Factor*) semi-supervisionado para detectar anomalias a partir de dados de tráfego espaço temporais (ST - *Spatial Temporal*), como por exemplo, identificação de acidentes e congestionamentos. Depois de aplicado a Análise de Componentes Principais (PCA - *Principal Component Analysis*) nos sinais de trânsito ST, com a finalidade de remover dados referentes a ruído, a implementação do LOF, com número de vizinho definido como 7, teve 93,5% de sucesso, utilizando como métrica de avaliação a Taxa de Sucesso de Detecção (DSR - *Detection Success Rate*), em relação a identificação de anomalias (MA, NGAN, LIU, 2016).

Zhao et al. propuseram um método de detecção de anomalias semi-supervisionado em relação às medidas de concentração de glicose, utilizando o algoritmo Ponto Central Independente ÓPTICO (ICP-OPTICS - *Independent Central Point OPTICS*). Para demonstrar a eficácia do método, foram utilizados o conjunto de dados reais e o simulador UVa/Padova, comprovando o sucesso em identificar pontos de anomalia em relação a concentração de glicose no sangue como também servindo como referencial para outras aplicações (ZHAO et al., 2022).

Neste artigo, serão comparados dois modelos de aprendizado de máquina do tipo semi-supervisionado com a finalidade de identificar condições normais e anormais do ISHM feito em um sistema de tubulação de baixo custo. Os modelos escolhidos são, Fator de Anomalia Local (LOF - *Local Outlier Factor*) (BREUNIG et al., 2000) e Máquina de Vetores de Suporte (SVM - *Support Vector Machine*) (CHANG, LIN, 2011). Para realizar esta comparação foram utilizadas as métricas de avaliação, Matriz de Confusão e curva Característica de Operação do Receptor (ROC - *Receiver Operating Characteristic*).

Este artigo foi dividido da seguinte maneira. Na seção 2, serão apresentados o referencial teórico sobre ISHM, das técnicas de aprendizado de máquina e finalizando com uma descrição sobre como foi conduzido o experimento. Na seção 3, serão abordadas as maneiras como foram feitas as implementações das técnicas bem como um comparativo entre as métricas de avaliação. Na seção 4, constará as considerações finais com toda a abordagem feita neste trabalho.

2. METODOLOGIA

A técnica ISHM utiliza pastilhas piezoelétricos ligadas a uma estrutura principal, desse modo, faz proveito de suas características de detecção e atuação para identificar danos. Com isto, é feito o acompanhamento em alterações de rigidez, absorção de energia e massa da estrutura, baseando se na propriedade de interação eletromecânica de estrutura inteligente (GIURGIUTIU, ZAGRAI, 2005).

A metodologia de aplicação deste método torna possível prevenir situações, como, por exemplo, manutenção desnecessária, elevados custos e/ou danos, por meio da implementação de um sistema de vigilância, reduzindo a necessidade de intervenções não essenciais e incrementando a segurança. De início, em termos de princípio direto, ocorre a geração de um potencial elétrico nos terminais devido à tensão mecânica que é exercida

sobre o transdutor. Tratando se do princípio inverso, sucede a produção de uma tensão mecânica à medida que se aplica uma tensão elétrica nos terminais do transdutor. Equações destes dois princípios podem ser encontrados em (MUGANDA et al., 2022).

Em 1994, Liang, Sun e Rogers publicaram uma pesquisa na qual foi apresentado um modelo capaz de reconhecer o procedimento de medição da impedância eletromecânica para um grau de liberdade específico. Foi feito, então, a fim de realizar os cálculos, a combinação da função de impedância mecânica de PZT com a impedância estrutural do sistema, a fim de se obter uma função denominada de função de admitância. Esta função pode ser encontrada em (LIANG, SUN, ROGERS 1994).

Apesar de que as assinaturas de impedância forneçam informações específicas relacionadas aos danos de estruturas, ao comparar a assinatura de uma estrutura íntegra com a assinatura posterior a um evento, é possível observar, qualitativamente, as mudanças nas características estruturais (rigidez local, massa e amortecimento), indicando a ocorrência de uma falha. A fim de se realizar um procedimento de modo quantitativo, faz se o uso de métricas de danos na qual, o mais usual, se trata de um valor escalar que fornece informações quantitativas, sendo este o índice, Raiz do Desvio Quadrático Médio (RMSD - *Root Mean Square Deviation*), que pode ser encontrada em (GIURGIUTIU, ZAGRAI, 2005).

Algoritmos de aprendizado semi-supervisionado consistem na combinação de técnicas do tipo supervisionado e não-supervisionado onde o banco de dados principal é dividido em dois subconjuntos, sendo o primeiro contendo dados com rótulos conhecidos e o segundo com rótulos desconhecidos (GÉRON, 2019; HAYKIN, 2009).

Para este artigo, dois modelos de detecção de anomalias serão comparados, sendo eles, LOF e SVM com o objetivo de identificar as classes vigentes em relação ao ISHM aplicado a um sistema de tubulação de baixo custo. Originalmente, estas técnicas de aprendizado de máquina são do tipo não-supervisionado e supervisionado, respectivamente. Mas para este artigo, algumas modificações na implementação foram feitas para transformar estes métodos em semi-supervisionado.

O algoritmo LOF, na sua forma original, leva em consideração apenas a vizinhança restrita de cada objeto. Para a maioria dos objetos em um agrupamento, o valor do LOF é aproximadamente igual a 1, enquanto outros objetos têm aplicados limites inferiores e superiores que destacam sua natureza local. O cálculo do LOF de cada objeto é baseado no número de vizinhos mais próximos na vizinhança do objeto, sendo esse número utilizado

para identificar possíveis anomalias locais. Conforme são realizadas as etapas apresentadas, procede-se ao cálculo da Densidade de Acessibilidade Local (LRD - *Local Reachability Density*). Nesse cálculo, são levados em conta dois parâmetros que determinam a densidade: o primeiro é o número mínimo de objetos e o segundo é o volume (BREUNIG et al., 2000).

Por outro lado, o algoritmo SVM, na sua forma original, utiliza-se do mapeamento dos dados de entrada em relação a criação de um espaço de características de alta dimensão, empregando uma função *Kernel* com o intuito de construir uma superfície de decisão linear neste espaço. Com esta superfície implementada, a classificação das classes que envolvem os dados de teste é identificada. A escolha da função *Kernel* varia de acordo com as condições que se encontra os dados de entrada, podendo ser uma função linear, polinomial, radial ou sigmoide (CORTES, VAPNIK, 1995; HEARST, 1998).

No experimento utilizado como base para este artigo, foi construído uma bancada de testes com propósito de emular os mecanismos de dano observados em tubulações industriais reais, apresentado pela Figura 1 (GARCIA et al., 2021).



Figura 1 - Bancada finalizada (Adaptado: GARCIA et al., 2021).

Neste contexto, buscou-se a aplicação da técnica de impedância eletromecânica para monitorar o comportamento dessa estrutura. Para esta finalidade, foi utilizado a placa Analog Devices, modelo EVAL-AD5933EBZ, apresentado pela Figura 2 (GARCIA et al., 2021).



Figura 2 - Analog Devices EVAL-AD5933EBZ (Adaptado: GARCIA et al., 2021).

No início do experimento, foram coletadas as amostras do estado natural, sem nenhum dano aparente. Em seguida, foram colhidas amostras para cada tipo de dano específico, sendo estes, encrustamento de menor magnitude provocado por um filtro, mostrado na Figura 3, entupimento utilizando massa epóxi, mostrado na Figura 4 e dano provocado por riscados, mostrado na Figura 5.



Figura 3 - Filtro (Adaptado: GARCIA et al., 2021).



Figura 4 - Massa epóxi (Adaptado: GARCIA et al., 2021).

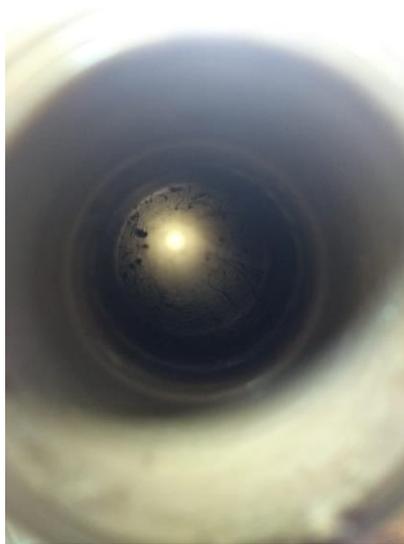


Figura 5 - Riscados (Adaptado: GARCIA et al., 2021).

Essa abordagem permitiu obter um conjunto completo de dados para cada cenário e faixa de frequência, possibilitando a avaliação do comportamento da tubulação em relação aos diferentes danos aplicados. Com isto, após a captação dos sinais de impedância eletromecânica, necessitou-se normalizar os sinais, devido a variação da temperatura, para que possam ser usados de modo comparativo de cada posição entre si (GARCIA et al., 2021). A Figura 6 apresenta o comportamento dos sinais de impedância eletromecânica em relação a frequência.

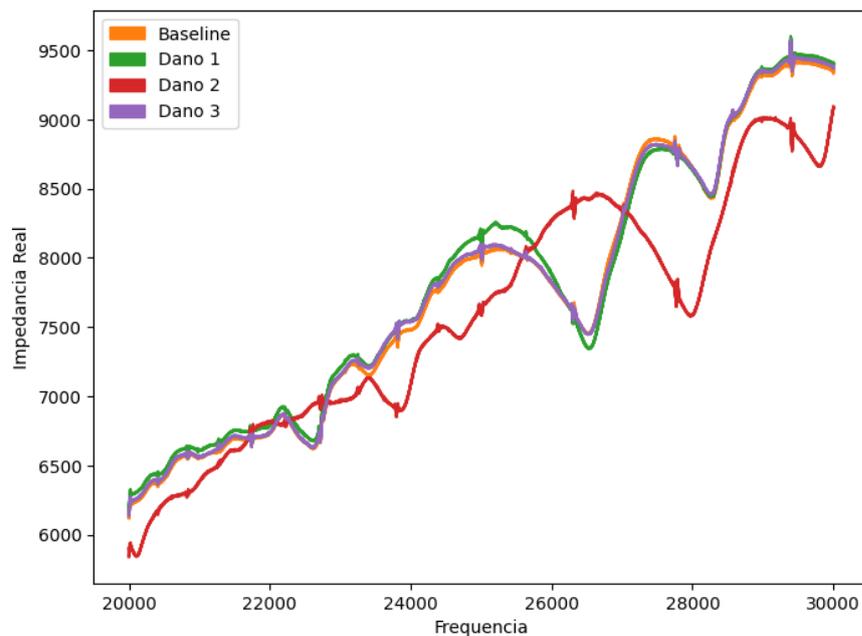


Figura 6 - Conjunto de sinais normalizados (Adaptado: GARCIA et al., 2021).

3. RESULTADOS E DISCUSSÃO

Neste estudo de caso, a implementação dos algoritmos LOF e SVM, bem como as métricas de avaliação Matriz de confusão e curva ROC foram realizadas na linguagem *Python* utilizando a biblioteca *Scikit-learn* (PEDREGOSA et al., 2011). Iniciando pela aplicação do LOF, cinco assinaturas de impedância do *baseline* foram utilizadas no treinamento com os parâmetros, $n_neighbors=4$ e $novelty=True$. Originalmente, este método é do tipo não-supervisionado, ou seja, os rótulos não são conhecidos, mas sim agrupados quanto a suas semelhanças. Alterando o parâmetro 'novidade' para verdadeiro, ou seja, $novelty=True$, a implementação deixa de ser aprendizado não-supervisionada e passa a ser aprendizado semi-supervisionado do tipo detecção de novidades, também chamado de aprendizado semi-não-supervisionado, onde os dados de treinamento são todos agrupados em uma classe, sendo este o estado natural da estrutura. Caso, na etapa do treinamento, ocorra uma mistura entre classes diferentes, os resultados estarão comprometidos. Com a obtenção deste rótulo, foi utilizado um banco de dados contendo, cinco pontos do estado natural e dez pontos dos três tipos de danos, totalizando trinta e cinco assinaturas de impedância para testar o modelo. A Figura 7 apresenta o resultado da aplicação da Matriz de confusão entre os rótulos verdadeiros e os que foram previstos.

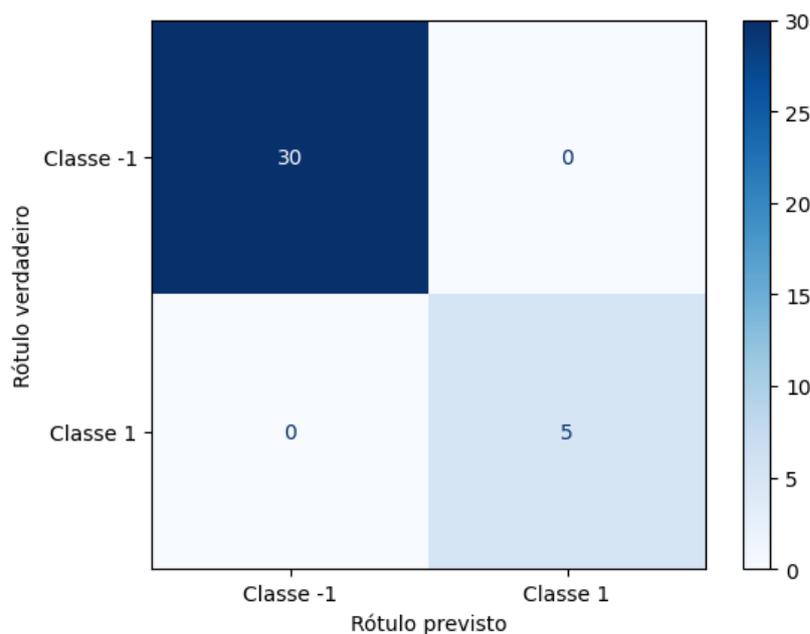


Figura 7 - Matriz de Confusão (LOF).

De acordo com a Figura 7, o modelo identificou trinta casos de anomalia (Classe -1) e cinco casos do estado natural (Classe 1) de maneira correta, portanto, todas as assinaturas de impedância foram identificadas corretamente. A Figura 8 apresenta o gráfico da curva ROC bem como o valor de AUC.

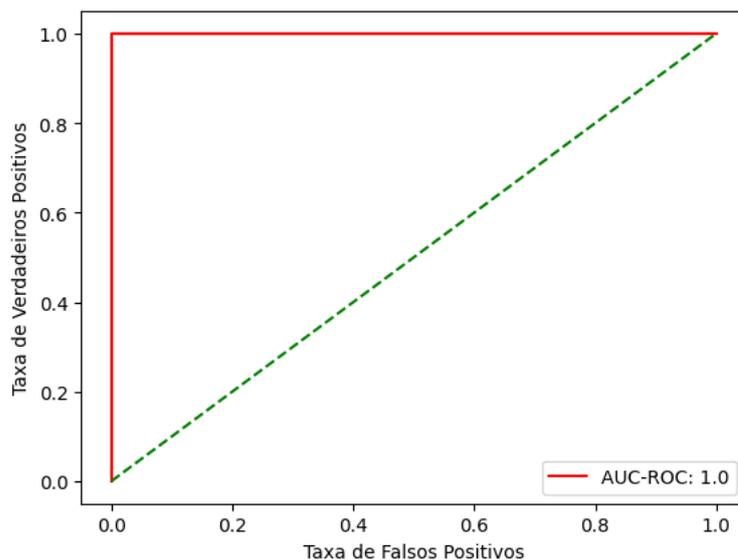


Figura 8 - Curva ROC (LOF).

Visto que o valor correspondente a AUC-ROC varia entre 0.0 e 1.0, sendo que, números próximos de 0.0 significa que o modelo não foi capaz de distinguir entre as duas classes e números próximos a 1.0 denota que o modelo teve um bom desempenho nas classificações, a aplicação desta métrica resultou no valor de 1.0, ou seja, o algoritmo LOF semi-não-supervisionado teve um desempenho ótimo na classificação das assinaturas de impedância do estado natural e dos danos no ISHM feito no sistema de tubulação.

Na aplicação do método SVM semi-supervisionado, o treinamento foi feito utilizando dois bancos de dados. O primeiro, é acompanhado junto com os rótulos, dispondo de treze assinaturas de impedância, sendo elas, cinco pontos do estado natural e oito pontos de dano. O segundo, contém treze pontos não rótulos, totalizando vinte e seis assinaturas. Inicialmente, é necessário criar a base do classificador utilizando SVM com os parâmetros, *kernel='linear'* e *probability=True*. Em seguida, utilizando a base criada anteriormente, um classificador semi-supervisionado foi criado pela função *SelfTrainingClassifier*. Esta função tem como objetivo transformar um classificador supervisionado em semi-supervisionado por meio de pseudo-rótulos para os dados não rotulados (PEDREGOSA et al., 2011). Por fim, foram utilizados os dois bancos de dados agrupados para realizar o treinamento, considerando os rótulos 0 (estado natural), 1 (dano) e -1 (pseudo-rótulo). Para testar o

modelo, foi utilizado um banco de dados contendo quatorze pontos de impedância, sendo, três assinaturas do estado natural e onze de dano. A Figura 9 apresenta a aplicação da Matriz de confusão para validar este método.

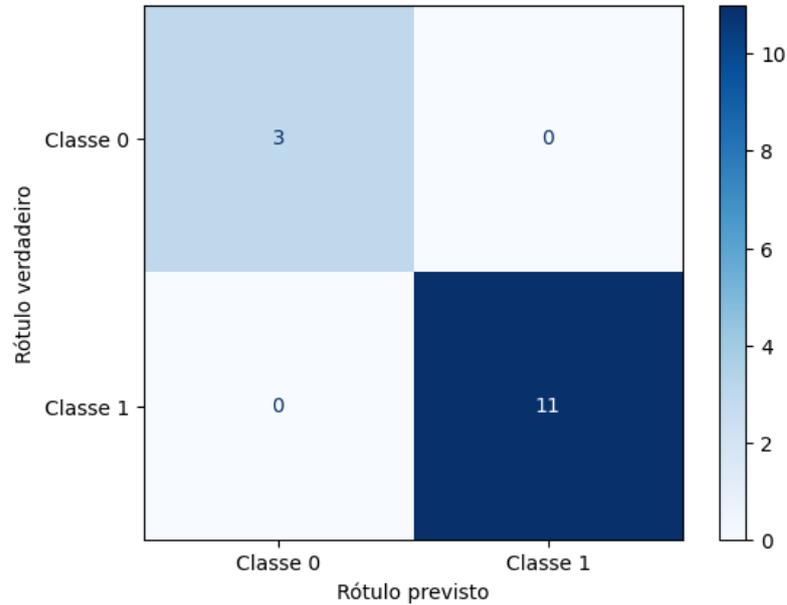


Figura 9 - Matriz de Confusão (SVM).

De acordo com a Figura 9, o modelo SVM semi-supervisionado identificou corretamente todas as assinaturas de impedância, sendo elas, três pontos da Classe 0 e onze da Classe 1. A Figura 10 apresenta o gráfico da curva ROC junto com o seu respectivo valor de AUC.

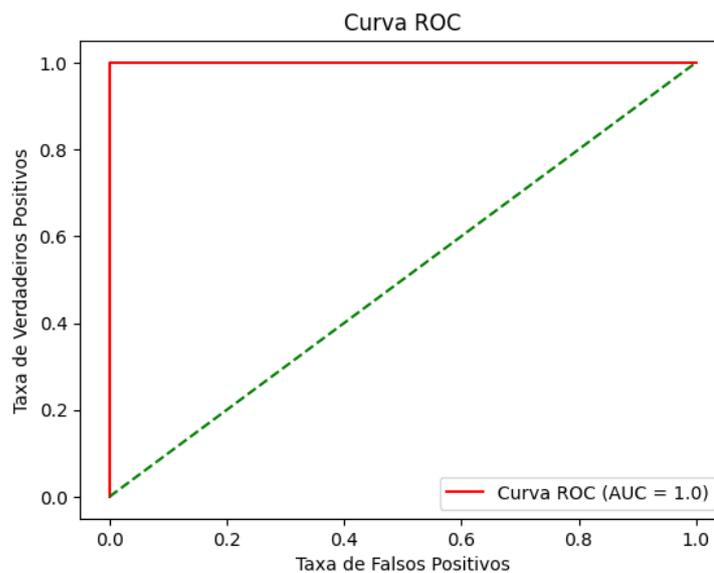


Figura 10 - Curva ROC (SVM).

Como pode ser observado na Figura 10, o valor de AUC-ROC é 1.0, ou seja, este método alcançou desempenho ótimo em relação a identificação das classes divergentes. Em função dos altos níveis de desempenho dos métodos, LOF semi-não-supervisionado e SVM semi-supervisionado, foi demonstrado, com o auxílio das métricas, Matriz de confusão e curva ROC, a alta eficiência das técnicas de detecção de anomalias baseadas em aprendizados semi-não-supervisionado e semi-supervisionado.

5. CONSIDERAÇÕES FINAIS

Neste trabalho foram comparadas duas vertentes de aprendizado de máquina, um semi-não-supervisionado, representado pelo método LOF (Originalmente não-supervisionado), e o outro semi-supervisionado, representado pelo método SVM (Originalmente supervisionado), com o objetivo de identificar de maneira correta, as condições normal e anormal de dados de impedância coletados a partir do monitoramento de integridade feito em um sistema de tubulação de baixo custo.

Um problema que vale ser ressaltado é na escolha dos dados de treinamento no algoritmo LOF, visto que, nesta etapa o método identifica todos os dados como sendo de uma única classe. Devido a isto, caso o conjunto seja composto por dados normais e anormais, os resultados estarão comprometidos. Já no algoritmo SVM, a escolha da função *Kernel* depende do comportamento dos sinais do banco de dados. No estudo de caso, a função *linear* foi escolhida devido a simplicidade deste comportamento, que pode ser observado na Figura 6.

Foram utilizadas como métricas de avaliação duas técnicas específicas, Matriz de Confusão e Curva ROC, no qual, por meio dos dados resultantes, houve sucesso em ambos os métodos, de forma que os dois modelos tiveram a taxa de desempenho máxima, ou seja, identificaram de maneira correta todos os pontos de assinatura.

A exatidão dos resultados apresentados pelas métricas de avaliação podem ter ocorrido devido a quantidade e simplicidade dos pontos de assinatura, que podem ser observados nas Figuras 7 e 9, onde um número muito baixo de dados de treinamento e teste foram disponibilizados. Mais testes são necessários utilizando bancos de dados maiores e mais complexos, com a finalidade de validar estas técnicas em detrimento de possíveis vieses que ocorrem devido ao baixo número de dados no conjunto de entrada.

AGRADECIMENTOS

O primeiro e o segundo autor agradecem a CAPES pela bolsa de mestrado recebida durante o desenvolvimento deste trabalho.

REFERÊNCIAS

BOUKERCHE, Azzedine; ZHENG, Lining; ALFANDI, Omar. Outlier detection: Methods, models, and classification. **ACM Computing Surveys (CSUR)**, v. 53, n. 3, p. 1-37, 2020.

BREUNIG, Markus M. et al. LOF: identifying density-based local outliers. In: **Proceedings of the 2000 ACM SIGMOD international conference on Management of data**. 2000. p. 93-104.

CHANDOLA, Varun; BANERJEE, Arindam; KUMAR, Vipin. Anomaly detection: A survey. **ACM computing surveys (CSUR)**, v. 41, n. 3, p. 1-58, 2009.

CHANG, Chih-Chung; LIN, Chih-Jen. LIBSVM: a library for support vector machines. **ACM transactions on intelligent systems and technology (TIST)**, v. 2, n. 3, p. 1-27, 2011.

CORTES, Corinna; VAPNIK, Vladimir. Support-vector networks. **Machine learning**, v. 20, p. 273-297, 1995.

GARCIA, Renan et al. Otimização de faixa de frequência dos sinais de impedância adquiridos no monitoramento da integridade estrutural em tubulações industriais. **Revista Cereus**, v. 13, n. 3, p. 127-141, 2021.

GÉRON, Aurélien. **Mãos à obra: aprendizado de máquina com Scikit-Learn & TensorFlow**. Alta Books, 2019.

GIURGIUTIU, Victor; ZAGRAI, Andrei. Damage detection in thin plates and aerospace structures with the electro-mechanical impedance method. **Structural Health Monitoring**, v. 4, n. 2, p. 99-118, 2005.

HAYKIN, Simon. **Neural networks and learning machines, 3/E**. Pearson Education India, 2009.

HEARST, Marti A. et al. Support vector machines. **IEEE Intelligent System and their applications**, v. 13, n. 4, p. 18-28, 1998.

HUSSAIN, Bilal; DU, Qinghe; REN, Pinyi. Semi-supervised learning based big data-driven anomaly detection in mobile wireless networks. **China Communications**, v. 15, n. 4, p. 41-57, 2018.

LIANG, Chen; SUN, Fan Ping; ROGERS, Craig A. Coupled electro-mechanical analysis of adaptive material systems-determination of the actuator power consumption and system energy transfer. **Journal of intelligent material systems and structures**, v. 5, n. 1, p. 12-20, 1994.

MA, Mathew X.; NGAN, Henry YT; LIU, Wei. Density-based outlier detection by local outlier

factor on largescale traffic data. **Electronic Imaging**, v. 28, p. 1-4, 2016.

MUGANDA, James M. et al. Influence function measurement technique using the direct and indirect piezoelectric effect for surface shape control in adaptive systems. **IEEE Transactions on Automation Science and Engineering**, v. 19, n. 2, p. 994-1002, 2021.

PEDREGOSA, Fabian et al. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825-2830, 2011.

WANG, Hongzhi; BAH, Mohamed Jaward; HAMMAD, Mohamed. Progress in outlier detection techniques: A survey. **IEEE Access**, v. 7, p. 107964-108000, 2019.

WANG, Xinlin; YANG, Insoon; AHN, Sung-Hoon. Sample efficient home power anomaly detection in real time using semi-supervised learning. **IEEE Access**, v. 7, p. 139712-139725, 2019.

ZEMMAL, Nawel et al. Adaptive semi supervised support vector machine semi supervised learning with features cooperation for breast cancer classification. **Journal of Medical Imaging and Health Informatics**, v. 6, n. 1, p. 53-62, 2016.

ZHAO, Yuhang et al. An independent central point OPTICS clustering algorithm for semi-supervised outlier detection of continuous glucose measurements. **Biomedical Signal Processing and Control**, v. 71, p. 103196, 2022.