



Editais FACEPE 01/2023
Programa Institucional de Bolsas de Iniciação Científica
PIBIC/FACEPE - 2023

RELATÓRIO FINAL

Projeto:
DESENVOLVIMENTO DE MODELOS EXPLICÁVEIS DE DETECÇÃO DE ANOMALIAS EM SÉRIES
TEMPORAIS BASEADOS NO PROJETO DE CARACTERÍSTICAS ORIENTADAS À APLICAÇÃO

Instituição Proponente e Executora:
Universidade Católica de Pernambuco

Aluno:
Gabriel Henrique Torres Santos

Curso:
Ciências da Computação

Coordenador:
Rodrigo de Paula Monteiro - rodrigo.paula@unicap.br

Período de vigência:
01/09/2023 a 31/07/2024

Número do processo:
BIC-1509-1.03/23

Período de realização:
2023.2-2024.1

Julho/2024

ABSTRACT:

A detecção de anomalias em séries temporais é importante para diversas aplicações, como segurança de rede, detecção de fraude e manutenção preditiva de equipamentos industriais. Este estudo propõe um modelo para a identificação de anomalias em séries temporais de sinais sonoros de dispositivos industriais, utilizando características derivadas de Coeficientes Cepstrais da Frequência de Mel (MFCCs). Especificamente, foram utilizadas features como Magnitude Mean Feature MFCC, Magnitude STD Feature MFCC e Correlation Feature MFCC, que capturam informações detalhadas sobre a correlação dos sinais de áudio. Diferentemente das abordagens baseadas em aprendizado profundo, que apresentam limitações em interpretabilidade e custo computacional, a técnica proposta oferece uma solução computacionalmente mais eficiente e interpretável. Os resultados indicam que o uso dessas características projetadas pode reduzir significativamente o custo computacional e melhorar a confiabilidade da detecção de anomalias.

PALAVRAS CHAVES: *MFCCs; Detecção de anomalias; Séries temporais; Sinais sonoros; Dispositivos industriais.*

1. INTRODUÇÃO:

A detecção de anomalias em séries temporais tem se tornado uma área de grande interesse em diversas indústrias, como saúde, finanças, tecnologia e fábricas com maquinário pesado. Séries temporais são conjuntos de dados coletados sequencialmente ao longo do tempo e são essenciais para monitorar e analisar tendências, padrões e comportamentos irregulares em sistemas dinâmicos. Em setores como saúde, industrial e tecnologia, a capacidade de identificar anomalias em séries temporais pode ajudar a detectar falhas em equipamentos, monitorar condições de saúde de pacientes e garantir a segurança dos sistemas [1].

No contexto industrial, as séries temporais são amplamente utilizadas para monitorar processos e operações em tempo real. Por exemplo, em uma linha de produção, sensores podem coletar dados continuamente sobre temperatura, pressão e velocidade [2]. Análises dessas séries temporais podem revelar desvios inesperados nesses parâmetros, que podem indicar problemas emergentes, permitindo que ações corretivas sejam tomadas antes que ocorram falhas. Dessa forma, a capacidade de monitorar essas variáveis em tempo real e responder rapidamente às anomalias é vital para a eficiência e segurança das operações industriais.

Apesar de sua importância, a detecção de anomalias em séries temporais enfrenta o desafio da falta de modelos explicativos robustos. Muitas abordagens atuais, baseadas em técnicas de aprendizado de máquina e inteligência artificial, podem detectar anomalias com alta precisão, mas normalmente carecem de transparência em seus processos de decisão. Isso resulta em uma "caixa preta", onde os usuários não conseguem compreender como e por que uma determinada anomalia foi detectada. Essa ausência de explicabilidade pode dificultar a confiança e a aceitação dessas tecnologias em ambientes críticos.

O estudo propõe uma abordagem inovadora para a detecção de anomalias em séries temporais de sinais sonoros de dispositivos industriais. Utilizando características derivadas de Coeficientes Cepstrais da Frequência de Mel (MFCCs), a pesquisa busca fornecer uma solução computacionalmente eficiente e interpretável. A contribuição esperada deste estudo é oferecer uma metodologia robusta que melhore a confiabilidade na detecção de anomalias, reduzindo o custo computacional e permitindo uma melhor compreensão do comportamento dos modelos aplicados em ambientes industriais.

2. OBJETIVO:

2.1 Objetivo geral:

Desenvolver e aprimorar modelos de detecção de anomalias em séries temporais, focando na melhoria da explicabilidade dos resultados, de modo a facilitar a interpretação e a tomada de decisão em ambientes industriais.

2.2 Objetivos específicos:

- Realizar uma revisão da literatura sobre técnicas de detecção de anomalias em séries temporais.
- Coletar conjuntos de dados reais de ambientes industriais.
- Treinar os modelos de detecção de anomalias utilizando os dados pré-processados.
- Validar a eficácia dos modelos através de métricas apropriadas, como acurácia.

3. METODOLOGIA

Para alcançar os objetivos apresentados neste projeto, o desenvolvimento do trabalho foi baseado no uso de metodologias ágeis. Isso envolveu entregas frequentes, interações semanais com o orientador e com o grupo de pesquisa, que foram essenciais para revisar o progresso, discutir desafios e ajustar as abordagens conforme necessário, adaptabilidade a mudanças e inovações, e melhorias contínuas no processo.

3.1. Coeficientes Cepstrais da Frequência de Mel (MFCC)

Usando como base do estudo o artigo [3], o MFCC é uma técnica de extração de características amplamente utilizada no processamento de sinais de áudio. Este método transforma um sinal de áudio em uma representação compacta que preserva informações essenciais sobre a estrutura espectral, facilitando a análise e reconhecimento de padrões. O estudo focou nos sons de máquinas industriais, utilizando o MFCC para monitoramento e detecção de falhas.

3.2. Dados de treino

Foram reunidos 588 dados para o treinamento do modelo e 240 para teste. Todos os dados de treinamento eram dados reais, permitindo ao modelo diferenciar melhor quando um dado estava fora do parâmetro esperado.

Os dados foram coletados do banco de dados MIMII (do inglês Malfunctioning Industrial Machine Investigation and Inspection) [4], utilizando oito microfones posicionados ao redor da máquina. Com esse conjunto de microfones, podemos avaliar tanto abordagens baseadas em um único canal quanto abordagens multicanais. No entanto, a abordagem utilizada neste estudo contou apenas com um canal. Os microfones foram mantidos a uma distância de 50 cm da máquina e gravaram segmentos sonoros de 10 segundos. O conjunto de dados contém oito canais separados para cada segmento. Além do som da máquina alvo, o ruído de fundo em várias fábricas foi continuamente gravado e posteriormente misturado com o som da máquina alvo para simular ambientes reais. O mesmo conjunto de microfones foi utilizado para gravar o ruído de fundo. Mais informações sobre o banco de dados podem ser encontradas no artigo [5].

3.3. Revisão da literatura

Foram realizadas leituras sobre o MFCC, uma técnica amplamente utilizada no processamento de sinais de áudio para extrair características relevantes. Foi mostrada sua aplicação em sistemas de transtornos de patologia de voz e análise da indústria, e os resultados alcançados

em termos de precisão e eficácia na análise de áudio [3]. Além disso, foram analisados três artigos de destaque que abordam a implementação de modelos preditivos para o monitoramento da qualidade de máquinas [5] [6]. Esses estudos destacam a importância de técnicas avançadas de análise de dados e aprendizado de máquina na manutenção preditiva, tanto no contexto industrial, onde a detecção precoce de falhas pode reduzir custos operacionais e evitar paradas não programadas quanto em sistemas de transporte, como o monitoramento da condição de vagões de metrô[4].

3.4. Problemas e relevância do estudo

A falta de explicabilidade em modelos preditivos pode acarretar em uma série de problemas que vão desde a desconfiança dos usuários até questões éticas e legais, como a falta de segurança e a perda da eficiência da máquina [2]. Quando um modelo não é capaz de explicar como chegou a determinada decisão ou previsão, a confiança dos usuários nos resultados é abalada, especialmente em áreas críticas como medicina e segurança. Outra preocupação é a possibilidade do modelo pegando padrões injustos e discriminação, já que modelos complexos podem aprender padrões discriminatórios nos dados sem que isso seja notado. Esse estudo propõe o uso de informações alternativas para a melhorar a confiança no modelo e para ter previsões mais precisas.

3.5 TREINAMENTO:

Neste tópico vai ser explicado o passo a passo do treinamento nos modelos preditivos, o pré-processamento dos dados e quais critérios foram usados para coletar os resultados finais. A Figura 1 apresenta todo este passo a passo.

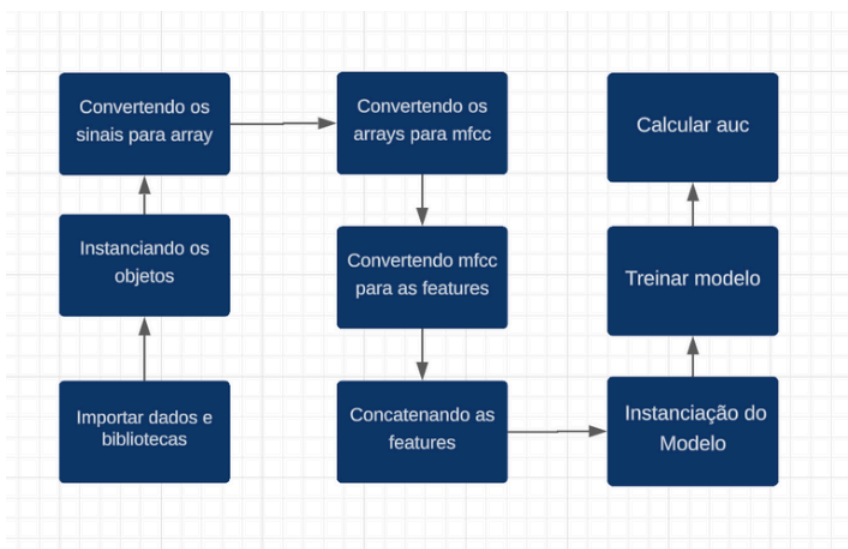


Figura 1 - ilustração do passo a passo do trabalho.

3.5.1. Pré-processamento

O pré-processamento dos dados seguiu os seguintes passos:

- Instanciação dos objetos: Inicializamos os objetos necessários para a manipulação e processamento dos sinais de áudio.
- Conversão dos sinais para arrays: Os sinais de áudio foram convertidos para arrays numéricos, facilitando seu tratamento computacional.

- Transformação em MFCCs: Esses arrays foram transformados em Coeficientes Cepstrais de Frequência Mel (MFCCs), uma representação compacta e eficiente da informação contida em sinais temporais, especialmente útil para problemas envolvendo áudios.
- Extração das features: A partir dos MFCCs, extraímos as features relevantes que o modelo de aprendizado de máquina irá utilizar.
- Concatenação das features: Finalmente, as features extraídas foram concatenadas, formando um conjunto de dados unificado para treinamento e avaliação do modelo.

3.5.2. Features

As features utilizadas neste estudo são o Magnitude Mean Feature MFCC, o Magnitude STD Feature MFCC e o Correlation Feature MFCC.

O Magnitude Mean Feature se refere ao valor médio de cada coeficiente MFCC ao longo do tempo. Em outras palavras, é a média dos valores dos coeficientes MFCC extraídos de um segmento de áudio. Essa feature é importante porque ajuda a capturar as características gerais do espectro de frequências do áudio.

O Magnitude STD Feature representa o desvio padrão dos coeficientes MFCC ao longo do tempo. O desvio padrão mede a variabilidade ou dispersão dos coeficientes MFCC em relação à média. Essa feature é crucial para entender a consistência ou a variação no comportamento dos coeficientes MFCC.

O Correlation Feature é utilizado para capturar a correlação entre diferentes coeficientes MFCC. A correlação mede o grau de dependência linear entre os coeficientes, indicando como os coeficientes variam em conjunto.

3.5.3. Modelos e treinamento

Os modelos utilizados neste estudo foram o One Class SVM, Isolation Forest, Local Outlier Factor e o GMM, mais informações sobre eles podem ser encontradas [3] [5]. Para comparar os modelos, seguimos um passo a passo que inclui o pré-processamento dos dados, a instanciação e o treinamento do modelo, para então calcular a acurácia. Esse processo foi repetido 20 vezes para cada configuração dos modelos, visando encontrar o melhor desempenho possível.

4. RESULTADOS:

Os resultados da obitidos do **MFCC** foram:

Tabela 1 - Média da acurácia obtido e o seu desvio padrão do MFCC

Parâmetros	AVERAGE acurácia
default	0,9173 (0,0417)

O MFCC foi testado apenas com os parâmetros padrão do GMM, foi o modelo que apresentou o melhor desempenho, com 91% de acurácia.

Os resultados da obitidos do **One Class SVM** foram:

Tabela 2 - Média da acurácia obtido e o seu desvio padrão do One Class SVM

parâmetros	AVERAGE acurácia
default	0.5137 (0.0333)
gamma='auto'	0.5830 (0.0335)
nu = 0.05	0.5174 (0.0432)
nu = 0.05 + gamma='auto'	0.5958 (0.0350)
nu = 0.05 + gamma='scale'	0.5187 (0.0450)

O One Class SVM apresentou o melhor desempenho com os parâmetros nu=0.05 (nu controla a proporção de outliers que o modelo deve permitir) e gamma=auto (gamma determina a influência de cada ponto de dados na decisão do modelo, com valores mais altos indicando que apenas pontos próximos afetam a decisão, enquanto valores mais baixos consideram pontos mais distantes), resultando em uma melhora de 8% dos parâmetros padrões, mas ainda assim, o modelo não apresentou resultados muito bons. As informações dos parâmetros do One Class SVM estão no site oficial do sklearn¹.

Os resultados da obitidos do **Isolation Forest** foram:

Tabela 3 - Média da acurácia obtido e o seu desvio padrão do Isolation Forest

Parâmetros	AVERAGE acurácia
default	0.8766 (0.0400)
bootstrap = True	0.8866 (0.0400)
max_features=0.05	0.8802 (0.0500)
max_features=0.5	0.8891 (0.0450)
n_estimators = 150	0.8969 (0.0400)
n_estimators = 200	0.8950 (0.0550)
n_estimators = 25	0.8649 (0.0650)
n_estimators = 50	0.8645 (0.0550)

O Isolation Forest teve um melhor resultado com a mudança no parâmetro n_estimators=150. Embora não tenha havido uma diferença muito grande do parâmetro padrão, todos os resultados foram bons, todos acima de 85%.As informações dos parâmetros do Isolation Forest estão no site oficial do sklearn².

¹ <https://scikit-learn.org/stable/modules/generated/sklearn.svm.OneClassSVM.html>

² <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.IsolationForest.html>

Os resultados da obitidos do **Local Outlier Factor** foram:

tabela 4 - Média da acurácia obtido e o seu desvio padrão do Local Outlier Factor

parâmetros	VERAGE acurácia
default	0.5579 (0.0400)
contamination = 0.05	0.5540 (0.0550)
contamination = 0.5	0.5579 (0.0450)
n_neighbors = 10	0.6084 (0.0600)
n_neighbors = 5	0.6509 (0.0640)
n_neighbors = 5 + contamination = 0.05	0.6399 (0,0350)
n_neighbors = 5 + contamination = 0.5	0.6423 (0,0450)

O Local Outlier Factor apresentou melhor desempenho com n_neighbors=5, mas, assim como o One Class SVM, não apresentou um desempenho muito bom. As informações dos parâmetros do Isolation Forest estão no site oficial do sklearn..As informações dos parâmetros do Local Outlier Factor estão no site oficial do sklearn.

A Figura 2 ilustra os resultados obtidos pelas melhores configurações de algoritmo. Os boxplots foram construídos com os resultados obtidos a partir de 20 modelos treinados, considerando cada algoritmo.

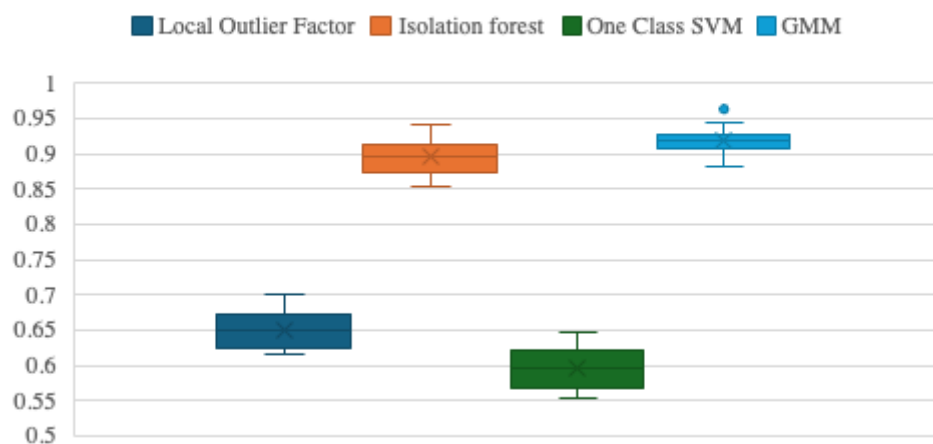


Figura 2: Boxplot para comparar os resultados obtidos pelos melhores modelos de cada algoritmo.

5. DISCUSSÃO:

A utilização das features Magnitude Mean Feature MFCC, Magnitude STD Feature MFCC e Correlation Feature MFCC proporciona uma abordagem robusta e eficaz para a análise de sinais de áudio no contexto de detecção de anomalias. Essas features são extremamente importantes porque

capturam informações essenciais sobre a natureza dos sinais de áudio, permitindo uma melhor compreensão e processamento pelos modelos de aprendizado de máquina.

Essas três features, quando usadas em conjunto, fornecem uma visão abrangente e detalhada dos sinais de áudio. Elas são capazes de capturar tanto as características globais quanto as variações internas e as inter-relações entre os componentes do sinal.

Além disso, ao utilizar essas features, conseguimos explorar as propriedades temporais e frequências dos sinais de áudio de maneira mais detalhada. Juntas, essas features criam um conjunto de dados robusto que melhora significativamente a capacidade dos modelos de aprendizado de máquina de distinguir entre comportamentos normais e anômalos, resultando em um desempenho superior na detecção de anomalias.

Por fim, essas features nos permitem entender melhor o comportamento dos modelos porque fornecem uma base rica e diversificada de dados sobre os sinais de áudio. Ao capturar diferentes aspectos dos sinais, essas features ajudam a identificar quais partes dos dados são mais influentes na decisão do modelo. Isso facilita a interpretação dos resultados, permitindo ajustar e otimizar os modelos de forma mais eficaz.

6. CONCLUSÃO:

O uso de características alternativas como correlação, variância e dispersão se mostrou essencial para a melhoria do processo de classificação, fornecendo informações valiosas que ajudam a diferenciar entre comportamentos normais e anômalos em máquinas industriais. A incorporação dessas características no desenvolvimento de modelos explicáveis de detecção de anomalias não só melhora o desempenho dos modelos, mas também fortalece a confiança nas previsões geradas, permitindo uma manutenção preditiva mais eficiente e segura nas indústrias.

Para trabalhos futuros, será testado o impacto de cada grupo de features individualmente para entender melhor suas contribuições específicas no processo de detecção de anomalias. Além disso, a implementação do SHAP (SHapley Additive exPlanations) pode ser explorada para melhorar ainda mais a interpretabilidade dos modelos [7]. O SHAP fornece explicações baseadas na teoria dos jogos que atribuem importância a cada feature, ajudando a identificar como cada característica influencia as previsões do modelo. Essas abordagens podem proporcionar uma compreensão mais profunda dos modelos e facilitar a sua aplicação prática em ambientes industriais.

7. Dificuldades encontradas:

Inicialmente, tive algumas dificuldades para entender o que eram séries temporais, pois nunca havia ouvido falar sobre esse conceito antes, o outro desafio foi entender como o uso de features poderia melhorar a explicabilidade do modelo. Com a leitura de artigos e as reuniões com o orientador, fui compreendendo a importância e a aplicação das séries temporais na análise de dados e entender o valor das características derivadas para a interpretação dos resultados do modelo.

No entanto, a parte mais complicada do projeto foi treinar os modelos e coletar os resultados. O processo de treinamento e avaliação dos modelos era bastante demorado, levando cerca de 30 minutos para encontrar a média da acurácia de um único modelo. Essa etapa exigiu muito tempo e concentração para evitar treinar certas combinações de parâmetros duas vezes, garantindo que todas as configurações fossem devidamente testadas e analisadas.

8. Atividades paralelas:

Além do desenvolvimento deste projeto, a única atividade paralela que foi realizada durante o período foi a participação de um grupo de programação competitiva. Essa atividade envolve a resolução de problemas algorítmicos em um ambiente de competição.

REFERÊNCIAS:

- [1] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.
- [2] Miguel E. P. Silva, Bruno Veloso and João Gama, "Predictive Maintenance, Adversarial Autoencoders and Explainability"
- [3] Z. K. Abdul and A. K. Al-Talabani, "Mel frequency cepstral coefficient and its applications: A review," *IEEE Access*, 2022
- [4] MIMII Dataset: Sound dataset for malfunctioning industrial machine investigation and inspection. arXiv preprint arXiv:1909.09347, 2019. APA
- [5] Hao Zhou, Ke Yu, Xuan Zhang, Guanlin Wu and Anis Yazidi, "Contrastive autoencoder for anomaly detection in multivariate time series"
- [6] Ya Su, Youjian Zhao, Chenhao Niu, Rong Liu, Wei Sun and Dan Pei, "Robust Anomaly Detection for Multivariate Time Series through Stochastic Recurrent Neural Network"
- [7] Scott M. Lundberg and Su-In Lee, "A Unified Approach to Interpreting Model Predictions"
- [8] Y. Chinniah, "Analysis and prevention of serious and fatal accidents related to moving parts of machinery," *Safety science*, vol. 75, pp. 163–173, 2015.
- [9] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, "Deep learning for anomaly detection: A review," *ACM Computing Surveys (CSUR)*, vol. 54, no. 2, pp. 1–38, 2021.
- [10] Q. Tian, J. Li, and H. Liu, "A method for guaranteeing wireless communication based on a combination of deep and shallow learning," *IEEE Access*, vol. 7, pp. 38 688–38 695, 2019.
- [11] M. Raza and U. Qayyum, "Classical and deep learning classifiers for anomaly detection," in 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST), IEEE, 2019, pp. 614–618.
- [12] R. de Paula Monteiro, M. C. Lozada, D. R. C. Mendieta, R. V. S. Loja, and C. J. A. Bastos Filho, "A hybrid prototype selection-based deep learning approach for anomaly detection in industrial machines," *Expert Systems with Applications*, p. 117528, 2022.
- [13] F. Dama and C. Sinoquet, "Time series analysis and modeling to forecast: a survey," arXiv preprint arXiv:2104.00164, 2021.