

Detecção de Anomalias em Robôs Industriais

1st Andre Saghaard
Centro de Informática - UFPE
Recife, Brazil
afs15@cin.ufpe.br

2nd Bianca Paes
Centro de Informática - UFPE
Recife, Brazil
bpas@cin.ufpe.br

3rd Caio Nascimento
Centro de Informática - UFPE
Recife, Brazil
cns@cin.ufpe.br

4th Rodrigo Gomes
Centro de Informática - UFPE
Recife, Brazil
rgar@cin.ufpe.br

Abstract—The growing adoption of industrial robots in manufacturing and automation environments has increased the demand for reliable monitoring and early fault detection systems. In industrial contexts, operational anomalies can result in physical damage to equipment, production downtime, risks to human safety, and high corrective maintenance costs. Regarding existing solutions, the literature ranges from analytical dynamic model based approaches, which compare actual behavior with physical simulations but rely on precise parameterization, to classical statistical methods and state estimation filters. More recently, data driven techniques based on Machine Learning and Deep Learning have gained prominence for their ability to model complex non linearities without the need for explicit physical modeling. This work investigates the problem of anomaly detection in industrial robots using time series data from Inertial Measurement Units (IMU), which capture information on acceleration, rotation, and magnetic fields across spatial axes. Different types of mechanical faults are considered, including collisions, mechanical overload, and external disturbances, which distinctly affect the robot's physical dynamics.

I. INTRODUÇÃO

A motivação para este estudo decorre da crescente adoção de robôs industriais em operações críticas, onde a execução de tarefas repetitivas em proximidade com operadores exige monitoramento constante. Neste contexto, a identificação automática de anomalias é essencial para detectar falhas mecânicas e interferências externas, garantindo a integridade do equipamento e a segurança do processo.

A análise baseia-se em dados de sensores inerciais (IMU), capazes de capturar a dinâmica do manipulador. O acelerômetro monitora vibrações e impactos abruptos (indicativos de colisões), enquanto o giroscópio registra a velocidade angular, permitindo identificar oscilações ou atrasos causados por cargas excessivas. O magnetômetro, mede a orientação do sistema em relação ao campo magnético terrestre, apresentando variações mais lentas e sendo sensível a interferências magnéticas inesperadas. Como o robô opera em ciclos sequenciais, os sinais normais apresentam periodicidade bem definida; desvios nesse padrão, como alterações de forma de onda ou picos de energia, constituem as assinaturas principais de condições anômalas.

O conjunto de dados utilizado provém de experimentos públicos (IEEE PerCom 2023) realizados com um braço

robótico UR3e em ambiente controlado. As amostras consistem em séries temporais coletadas a 10 Hz durante tarefas de *pick-and-place*, abrangendo tanto o comportamento normal quanto cenários de falhas simuladas, como colisões (base e braço) e variações de carga. Essa diversidade torna o *dataset* ideal para avaliar a eficácia de técnicas de análise exploratória, pré-processamento e aprendizado de máquina na caracterização de falhas industriais.

II. ANÁLISE DE DADOS E FEATURE ENGINEERING

A. Análise Exploratória dos Dados

1) *Análise Exploratória Estrutural*: A análise exploratória estrutural foi conduzida com o objetivo de compreender a organização e os tipos de atributos presentes no conjunto de dados. Para garantir a integridade da avaliação durante as etapas de *feature selection* e modelagem de forma a prevenir a tomada de decisão enviesada e o vazamento de dados, adotou-se uma estratégia de particionamento dos dados a priori. A divisão em subconjuntos de Treino, Validação e Teste foi realizada imediatamente após a coleta. O Conjunto de Validação, composto por 10% dos dados Normais e 50% dos dados de Falha (de cada cenário), foi utilizado para a Análise Exploratória de Dados (EDA) e demais etapas subsequentes.

O dataset é composto por séries temporais de sensores inerciais, contendo variáveis numéricas contínuas associadas às medições de acelerômetro, giroscópio e magnetômetro, nos eixos X, Y e Z; uma coluna temporal que representa o instante de coleta de cada amostra e uma variável indicativa da condição de operação do robô, utilizada como rótulo para diferenciar entre comportamento normal e anômalo. Particularmente, os dados anômalos são separados de acordo com o tipo de falha, sendo elas:

- Hitting Platform, que lida com colisão contra a plataforma;
- Hitting Arm, colisão contra o próprio braço robótico;
- Extra Weight, esforço mecânico dado por peso extra;
- Earthquake, vibração estrutural externa.

Cada observação corresponde a uma leitura dos sensores em um determinado instante de tempo, sendo os dados organizados de forma sequencial. A variável temporal é representada por valores inteiros de alta resolução, indicando

registros em escala de nanosegundos, reforçando o caráter temporal do problema. No entanto, com o objetivo de facilitar a interpretação e a análise de integridade temporal, os valores foram convertidos para milissegundos, mantendo a proporcionalidade entre as amostras e reduzindo a magnitude numérica dos registros.

Durante a análise estrutural, foi identificada a presença de uma coluna textual associada à identificação do sensor, que por apresentar valor constante em todas as observações e não agregar informação discriminativa ao problema de detecção de anomalias, essa variável foi removida do conjunto de dados. Adicionalmente, foi realizada a verificação de duplicidade nos registros temporais, não sendo identificadas amostras com timestamps repetidos, o que indica consistência na indexação temporal das observações.

2) *Informações básicas*: O conjunto de dados normais analisado é composto por 874.937 amostras e 11 atributos após a etapa inicial de pré-processamento. Todas as variáveis de sensores são numéricas contínuas do tipo *float64*, enquanto a variável alvo é representada por valores inteiros binários. O volume total de memória ocupado por esse dataset é de aproximadamente 73 MB.

Por outro lado, o conjunto de dados anômalos é composto por 49.185 amostras com os 11 atributos contínuos numéricos e a variável alvo binária. Com um volume total de memória de aproximadamente 10MB. Adicionalmente, o dataset possui um atributo *scenario* que indica o tipo de falha a qual esse registro é associado, seguindo a proporção destacada na tabela abaixo.

TABLE I
QUANTIDADE DE ANOMALIAS POR TIPO DE FALHA

Cenário	Quantidade	Proporção
Hitting Platform	14.967	0.3043
Hitting Arm	11.924	0.2424
Earthquake	11.409	0.2320
Extra Weight	10.885	0.2213

3) *Análise de Duplicatas, Valores Faltantes e Outliers*: A integridade dos dados foi avaliada sob três perspectivas: redundância de registros, continuidade temporal e presença de valores extremos.

a) *Duplicatas e Valores Nulos Explícitos*: Uma varredura inicial no conjunto de dados brutos não identificou linhas duplicadas ou valores nulos (*NaN*) explícitos nas leituras dos sensores, tanto nos dados normais quanto anômalos. O sistema de aquisição registrou continuamente as 9 variáveis do IMU (acelerômetros, giroscópios e magnetômetros) sem falhas de escrita evidentes.

b) *Análise de Continuidade Temporal*: Embora não houvesse *NaNs* no arquivo original, a análise do intervalo entre amostras revelou um problema crítico de integridade temporal. A frequência nominal de coleta, era de 10 Hz (com *Sampling Rate* de 100 ms). No entanto, observou-se um *Jitter* (desvio padrão do intervalo de tempo) relevante de aproximadamente 22 ms, com 429 *Gaps* Temporais significativos variando entre 2 ms e 342 ms, onde essa quantidade de perda de pacotes foi

considerada quando o valor do *Gap* foi maior que 2x a média, sendo categorizada como quebra de continuidade.

Para padronizar a frequência e facilitar o janelamento das séries temporais nas etapas seguintes, foi utilizada a técnica de *Resampling*, impondo uma grade temporal rígida de 100 ms, conforme detalhado na seção de Pré-Processamento dos Dados. Nesse contexto, a irregularidade original foi exposta sob a forma de *Gaps* temporais. No conjunto de dados Normal, o alinhamento gerou cerca de 44.766 lacunas onde não havia dados registrados no timestamp esperado.

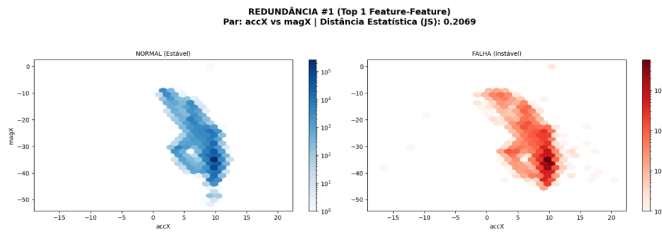
c) *Análise de Outliers*: A análise de distribuição, por meio de Boxplots e Histogramas, detectou uma quantidade massiva de *outliers* estatísticos, especialmente nos eixos do acelerômetro. Nos dados de colisão (*Hitting Arm*), a curtose do eixo Z atingiu valores extremos (> 100), com picos de aceleração de até 10g, muito superiores à faixa média normal de operação ($\pm 1g$). Entretanto, diferente de problemas de regressão clássicos, onde *outliers* são ruídos a serem removidos, neste projeto eles constituem o próprio **signal de interesse** (a falha mecânica).

4) *Análise Univariada*: Executar análises entre atributos e alvo de forma univariada com descrição estatística e visualizações de apoio.

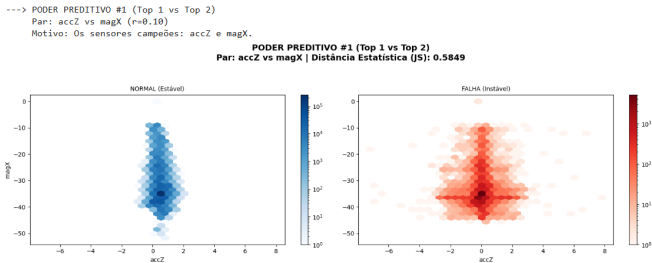
5) *Análise Bivariada*: A análise bivariada partiu da hipótese de que a ocorrência de anomalias afeta o estado físico do robô, questionou-se se as falhas seriam capazes de "quebrar" o acoplamento mecânico entre sensores redundantes ou se as alterações seriam mais perceptíveis em pares de baixa correlação. Para investigar essa premissa, foram definidos cenários analíticos estratégicos, cujo objetivo foi garantir que diferentes tipos de relação entre sensores fossem explicitamente investigados. Essa escolha é importante porque pares de variáveis com comportamentos distintos (alta correlação, correlação média ou baixa correlação) tendem a responder de maneira diferente à introdução de falhas, fornecendo evidências complementares sobre a dinâmica do sistema. A métrica utilizada para quantificar a divergência entre as distribuições do alvo normal e anômalo foi a distância de Jensen-Shannon.

Foram definidos seis cenários bivariados. Nos dois cenários de redundância (pares com maior correlação física), a importância reside em avaliar sensores fortemente acoplados. Nesses casos, observou-se que, mesmo na presença de falhas, o padrão conjunto se mantém relativamente estável. Quantitativamente, isso se traduziu nas menores distâncias estatísticas observadas: o par (*accx, accy*) apresentou uma distância JS de apenas 0.12, seguido pelo par (*accy, gyroy*) com 0.14. Esses baixos valores indicam que falhas nem sempre rompem relações altamente correlacionadas, sugerindo que pares muito redundantes carregam informação similar e tendem a variar de forma conjunta, mesmo em condições anômalas.

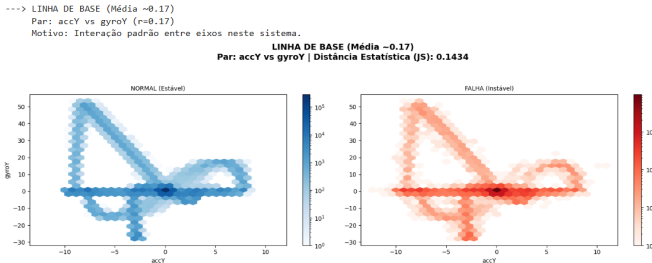
O cenário de linha de base, definido por um par com correlação próxima à média do sistema, foi incluído para representar um comportamento típico. O par (*accx, magx*) com uma distância JS de 0.20, ilustra essa referência. Esse valor intermediário permite avaliar como a resposta observada em



cenários extremos de fato se diferencia do comportamento esperado em uma interação padrão entre sensores.

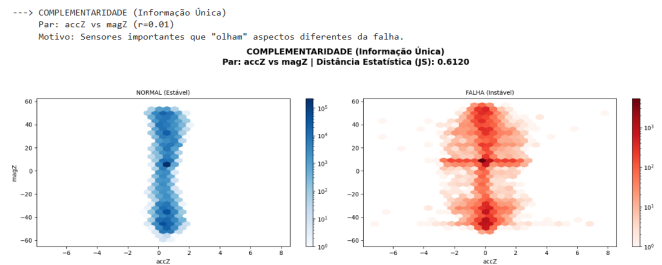


Já no cenário de complementaridade, composto por sensores pouco correlacionados entre si, observou-se a mudança mais pronunciada entre os dados normais e de falha. A baixa correlação implica que cada sensor responde a aspectos diferentes do fenômeno físico. Quando ocorre uma falha, essas respostas divergentes ampliam drasticamente a separação entre as distribuições conjuntas. Isso ficou evidente nos pares: (accz,magz) atingiu a maior distância registrada (0.61). Esse alto valor confirma que a combinação de variáveis complementares é a mais sensível para capturar a divergência estatística introduzida pelas anomalias.

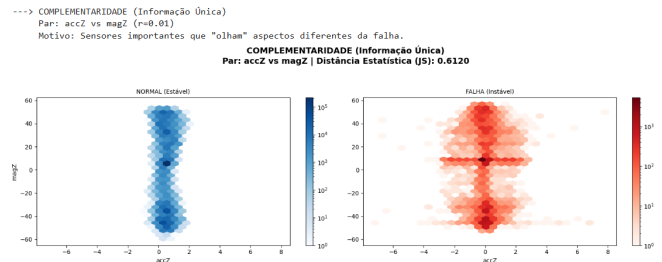


Por fim, os dois cenários de poder preditivo concentraram-se nos sensores mais correlacionados com o rótulo de falha. A importância desses cenários está em avaliar se sensores individualmente informativos também apresentam mudanças relevantes quando analisados em conjunto. Os resultados indicaram que a introdução de falhas altera a densidade conjunta, embora de forma menos extrema que na complementaridade, ou seja, o que importa não é unicamente a importância do sensor, mas também sua correlação para definir se o par vai se comportar de maneira mais parecida ou não.

Em suma, a análise validou a hipótese inicial: pares altamente correlacionados tendem a mascarar a anomalia, enquanto pares com baixa correlação evidenciam as alterações de forma mais clara.



6) Análise Multivariada: Na análise multivariada, todas as variáveis dos sensores foram consideradas simultaneamente por meio do PCA (Principal Component Analysis), após padronização. A projeção nos dois primeiros componentes revelou que as anomalias não são facilmente separáveis dos dados normais. As amostras de falha aparecem cercadas por pontos normais e distribuídas em diferentes regiões do espaço reduzido, indicando que não existe um único padrão geométrico simples que caracterize todas as falhas.

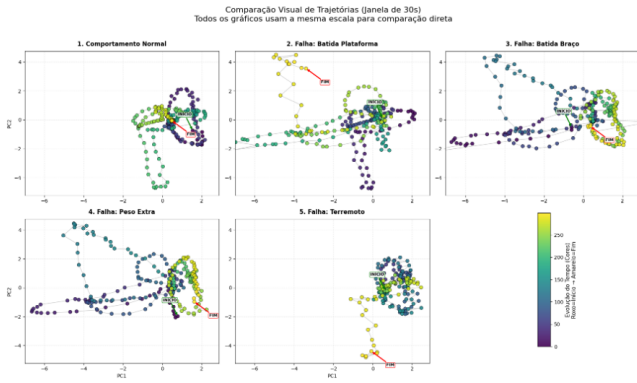


Apesar dessa sobreposição, a visualização evidenciou diferenças estruturais relevantes. Em diversas regiões do espaço PCA, observam-se concentrações de falhas com orientações e dispersões distintas, refletindo o impacto específico de cada tipo de anomalia sobre o sistema. Isso indica que, embora não haja separação linear clara, as falhas alteram a dinâmica multivariada de maneira consistente, produzindo padrões locais diferenciados.

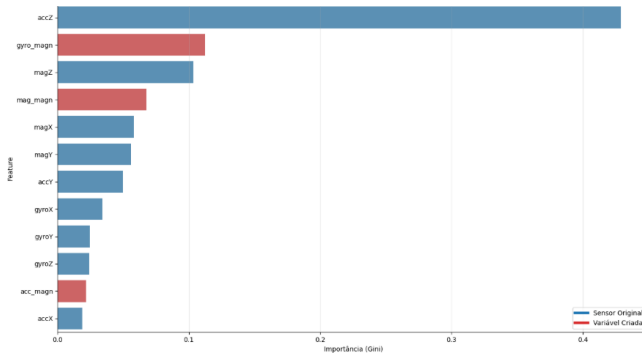
As setas do biplot permitiram interpretar a contribuição de cada sensor para essas diferenças, destacando quais variáveis influenciam mais fortemente determinadas direções do espaço PCA. Sensores com vetores mais extensos tiveram maior peso na organização dos dados, ajudando a explicar por que determinadas anomalias se manifestam em regiões específicas do plano.

A análise foi complementada pela visualização temporal de janelas de 30 segundos no espaço PCA, utilizando uma régua visual fixa para todos os cenários. Essa abordagem evidenciou que o comportamento normal tende a produzir trajetórias mais compactas e regulares, enquanto os cenários de falha apresentam trajetórias mais dispersas e irregulares, reforçando o impacto dinâmico das anomalias ao longo do tempo.

No que se refere às variáveis derivadas, foram calculadas magnitudes físicas da aceleração, do giroscópio e do magnetômetro, com o objetivo de condensar informações multieixo em medidas escalares fisicamente interpretáveis.



Testes exploratórios básicos, como a comparação de distribuições entre classes e análises de separação estatística, indicaram que essas variáveis capturam alterações relevantes introduzidas pelas falhas, mostrando potencial para uso na otimização dos modelos subsequentes, permitindo, por exemplo, estratégias de redução de dimensionalidade que mantenham a interpretabilidade física do fenômeno ao diminuir o número de variáveis de entrada.



Por fim, aplicou-se o K-Means para identificar grupos de interesse no espaço formado pelas variáveis originais. O uso de um método não supervisionado é adequado nesse contexto, pois permite investigar a estrutura intrínseca dos dados sem recorrer aos rótulos. Os resultados mostraram que, considerando apenas os dados brutos, não foi possível identificar um grupo claramente dominante em termos de concentração de falhas.

Cluster	Normal (%)	Falha (%)
0	96.601872	3.398128
1	96.528435	3.471565
2	96.646539	3.353461

B. Pré-processamento dos dados

Apresentar ações referentes ao pré-processamento de dados a exemplo de:

1) *Tratamento de Valores Faltantes*: Para corrigir os problemas de *timestamps* irregularmente espaçadas e amostras faltantes, sem descartar dados, optou-se pela **interpolação linear** dos valores faltantes gerados pelo *Resampling*. Essa abordagem preservou a tendência do movimento entre os

pontos conhecidos, restaurando a continuidade necessária para a extração de *features* de janela deslizante.

Para os conjuntos de dados anômalos, a etapa de regularização temporal com *Resampling* e Interpolação foi aplicada de forma iterativa e independente para cada cenário de falha, ao invés da aplicação direta sobre o conjunto anômalo completo. Essa abordagem foi necessária para prevenir as descontinuidades temporais, visto que os experimentos de falha foram registrados em contextos diferentes e com *Gaps* de minutos a horas entre si e a imposição de uma grade temporal contínua de 10Hz no dataset completo forçaria a criação de milhares de linhas vazias para preencher esses intervalos e a interpolação linear geraria dados sintéticos falsos, o que tornaria inviável alimentar o modelo com esses dados, mesmo com etapas subsequentes de pré-processamento.

2) *Tratamento de Outliers e Feature Scaling*: Conforme detalhado anteriormente, os *outliers* identificados representam uma parte importante do problema e representam medições verossímeis. Consequentemente, optou-se por **não removê-los**. No entanto, tendo em mente a fonte dos dados e a natureza dos sensores, foi necessário aplicar um filtro para reduzir o ruído dos sinais, sendo utilizado o Savitzky-Golay, que ajusta polinômios localmente, por sua capacidade de preservar melhor a forma e pico dos sinais que outros baseados em médias rolantes. Adicionalmente, para mitigar o impacto desses valores extremos dos sensores na normalização dos dados, substituiu-se o *StandardScaler* (sensível à média/desvio padrão) pelo **RobustScaler**, sendo este ajustado somente ao conjunto de treinamento dos dados normais e posteriormente usado para transformar os outros conjuntos. Este escalonador utiliza a mediana e o intervalo interquartil (IQR), garantindo que os picos de colisão permaneçam destacados na escala transformada, preservando a assinatura da anomalia para o modelo.

3) *Deteção e Tratamento de Duplicadas*: Não havia duplicadas no *dataset*, devido ao processo de geração deste [?].

4) *Encoding de Variáveis Categóricas*: Os dados não contém qualquer variável categórica, sendo desnecessárias estratégias de *encoding*.

C. Divisão dos Dados

Os dados normais, não-anômalos, conforme ilustrado em II, foram divididos em 60%, 20% e 20% entre os conjuntos de treinamento, validação e teste, respectivamente, enquanto foi dividido em 50% para a validação e 50% para o conjunto de testes. Em ambos os casos, a divisão respeitou a ordenação dos dados ao longo do tempo, visando evitar o vazamento de informações e a violação das relações temporais entre os pontos de dados.

Além disso, no caso dos dados anômalos, foi feita divisão estratificada, para garantir que cada classe de anomalia, que ocorreram cada uma independentemente e não concomitantemente, como os choques na plataforma e os choques no braço robótico, estivessem proporcionalmente representadas nos conjuntos de validação e teste.

TABLE II
DIVISÃO PERCENTUAL DOS CONJUNTOS DE DADOS)

Conjunto de Dados	Treino (%)	Validação (%)	Teste (%)
Normal	60,0	20,0	20,0
Anomalia	0,0	50,0	50,0

D. Feature Engineering

A estratégia de engenharia de atributos (*Feature Engineering*) adotada neste trabalho foi adaptada aos diferentes de tipos de modelos avalaiados. Processo de extração manual de atributos foi aplicado **exclusivamente** aos modelos probabilísticos e baseados em densidade (Isolation Forest, Gaussian Mixture Models).

Para as abordagens baseadas em Deep Learning, foram usados os dados brutos das séries temporais, submetidos apenas à criação de janelas, considerando a capacidade dessas arquiteturas de fazer (*representation learning*) e capturar dependências temporais complexas. Em contrapartida, os modelos clássicos carecem da habilidade nativa de interpretar a sequencialidade temporal de alta dimensão dos sensores brutos (90 dimensões por segundo de janela) tornando imperativa a transformação desses dados em vetores de estatísticas descritivas que sintetizem o comportamento dinâmico e espectral do robô.

Dessa forma, o *pipeline* de engenharia de atributos para os modelos clássicos foi feito em três etapas: Criação de janelas, extração e seleção, conforme detalhado a seguir.

1) *Criação de janelas*: Para a criação de janelas, foram utilizadas janelas móveis (*sliding windows*)??, que possuem uma intercessão ajustável com as janelas imediatamente vizinhas. Um algoritmo ilustrativo para a criação de janelas deslizantes esa descrito em 1.

Cada parte dos dados passou por esse processo isoladamente, e janelas incompletas foram descartadas para garantir a uniformidade das matrizes resultantes. Por exemplo, visto que a frequência de coleta de dados foi de 10Hz, uma janela de 40 amostras captura 4 segundos de funcionamento do braço robótico. Ademais, os parâmetros das janelas foram considerados como hiperparâmetros na etapa da modelagem, para que sejam escolhidos otimamente. Esta escolha é muito importante, porque o tamanho da janela depende muito dos modelos, e os eventos de interesse são muitos particulares do problema. No caso desse dataset, observamos que os eventos de interesse de fato exigiam uma análise um período de alguns segundos.

2) *Extração e seleção de features*: Nesta etapa, buscamos transformar os sinais dos sensores em uma representação mais informativa, bem como adaptá-la às limitações dos modelos que não usam Deep Learning usados neste trabalho, que não lidam bem com a grande quantidade de dimensões envolvida e com a alta correlação entre features. Esse processo se dividiu entre a extração e a seleção de features.

Inicialmente, fizemos a extração de features a nível de amostra. A partir dos sinais originais, criamos variáveis derivadas utilizando a interpretação física do problema e do

Algorithm 1 Segmentação em Janelas Deslizantes (Sliding Windows)

Require: D (Dados $N \times F$), L (Tamanho da Janela), O (Overlap)

Ensure: \mathcal{W} (Tensor de janelas $M \times L \times F$)

$S \leftarrow L - O$

$N \leftarrow \text{rows}(D)$

$\mathcal{W} \leftarrow \emptyset$

$i \leftarrow 0$

while $i + L \leq N$ **do**

$w \leftarrow D[i : i + L, :]$

$\mathcal{W} \leftarrow \mathcal{W} \cup \{w\}$

$i \leftarrow i + S$

end while

return \mathcal{W}

que cada variável representava, como a norma vetorial de cada sensor, que captura a magnitude total da mudança por ele registrada, unindo os eixos.

Em seguida, os sinais e seus atributos derivados foram segmentados em janelas como descrito a cima. Para cada janela, foram extraídas estatísticas agregadas calculadas exclusivamente com as *features* originais, agregando tendências e informações sobre o periodo inteiro de tempo representado pelo tamanho da janela.

- 1) **Média**: representa o valor médio do sinal ao longo da janela e está associada ao nível basal ou tendência local do movimento. Alterações na média podem indicar mudanças sistemáticas no comportamento do sistema, como inclinações persistentes ou deslocamentos contínuos.
- 2) **Desvio padrão**: quantifica a variabilidade do sinal dentro da janela, sendo fortemente associado à intensidade de vibrações.
- 3) **RMS (Root Mean Square)**: mede a energia total do sinal, combinando informações de magnitude e variabilidade. Diferentemente da média, o RMS é sensível tanto a oscilações quanto a impactos.
- 4) **Pico-a-pico (peak-to-peak)**: corresponde à diferença entre os valores máximo e mínimo do sinal na janela, também capturando grandes variações.
- 5) **Curtose**: caracteriza o grau de impulsividade do sinal, indicando a presença de picos abruptos e eventos raros de grande magnitude..
- 6) **Fator de crista (crest factor)**: definido como a razão entre o valor de pico e o RMS, fornece uma medida normalizada da severidade dos picos em relação à energia média do sinal. Essa *feature* é especialmente útil para diferenciar vibrações contínuas de impactos pontuais.
- 7) **Frequência dominante**: extraída a partir da transformada de Fourier da janela, indica a frequência com maior energia espectral. Essa informação permite capturar características dinâmicas do sistema, como ressonâncias, mudanças de regime de operação ou vibrações induzidas por cargas externas.

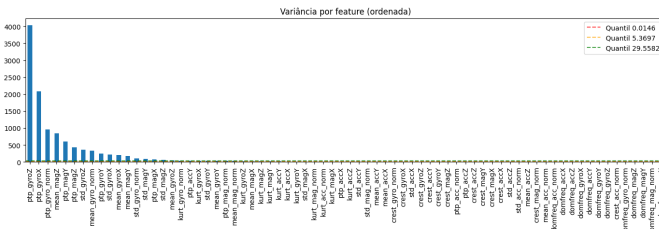
Essas *features* sintetizam muitas propriedades fundamentais do comportamento do sinal dentro da janela e constituem uma representação comparativamente compacta desse comportamento. Dessa forma, ao invés de termos um total de $features * tamanhoDaJanela$ dimensões, ou seja, 360, no caso dos nossos 9 sensores e uma janela de 40 amostras, obtemos apenas 72 dimensões.

No entanto, apesar de representar uma redução considerável da dimensionalidade, essas *features* ainda podem apresentar elevada redundância estatística. Em particular, muitas das estatísticas extraídas descrevem aspectos relacionados do sinal (por exemplo, desvio padrão, RMS e pico-a-pico), o que resulta em fortes correlações lineares entre atributos. A presença desse tipo de redundância pode prejudicar o desempenho de modelos probabilísticos e baseados em densidade, além de aumentar o custo computacional e dificultar a interpretação dos resultados.

Dessa forma, foi aplicada uma etapa explícita de seleção de *features*. Inicialmente, realizou-se uma análise de correlação linear entre os atributos extraídos, utilizando o coeficiente de correlação de Pearson. Para cada par de *features* cuja correlação absoluta excedia um limiar de 0,95, apenas uma delas foi mantida. Essa estratégia visa eliminar atributos altamente redundantes, preservando, ao mesmo tempo, a maior diversidade possível de informações relevantes. Após essa etapa, o número de *features* foi reduzido de 72 para 58.

Mesmo após a remoção de correlações elevadas, o espaço de atributos ainda permanecia relativamente grande e com possíveis dependências lineares de ordem superior. Assim, aplicou-se uma etapa adicional de redução de dimensionalidade por meio da Análise de Componentes Principais (PCA). O PCA foi ajustado exclusivamente sobre o conjunto de treinamento, após normalização robusta dos dados, e o número de componentes foi selecionado de forma a preservar pelo menos 95% da variância total dos dados. Como resultado, o espaço de atributos foi reduzido de 58 para 15 componentes principais, mantendo uma variância explicada acumulada de aproximadamente 95,28

Essa combinação de seleção por correlação e redução via PCA permitiu obter uma representação compacta, pouco redundante e estatisticamente bem condicionada dos dados, adequada às premissas dos modelos clássicos avaliados neste trabalho. Além disso, ao concentrar a maior parte da variabilidade dos sinais em um número reduzido de componentes, essa abordagem contribui para maior estabilidade numérica, melhor generalização e menor sensibilidade a ruído.



III. MODELAGEM

Três modelos foram escolhidos para esse projeto: Gaussian Mixture Model (GMM), Isolation Forest e Autoencoder Convolutacional. Todos os modelos condizem com o problema proposto: são próprios para aprendizado não-supervisionado, cada um possuindo uma abordagem distinta para tanto: clustering, isolamento e aprendizado profundo baseado em redes neurais, respectivamente, tornando-os apropriados para comparar a eficácia desses diferentes métodos para o dataset em questão. O processo de seleção de modelos utilizou como métrica principal a AUC. No entanto, devido ao uso de predições binárias em vez de escores contínuos, essa métrica não corresponde estritamente à AUROC tradicional, funcionando como uma medida de desempenho baseada em classificação após a aplicação de um limiar

A. Gaussian Mixture Model

1) *Conceitos Básicos:* O modelo de mistura gaussiana (gaussian mixture model, GMM) é um modelo probabilístico que assume que os dados são descritos por uma mistura de distribuições gaussianas.

Formalmente, um GMM modela a densidade de probabilidade de uma observação $\mathbf{x} \in R^d$ como uma soma ponderada de K componentes Gaussianos, dada por:

$$p(\mathbf{x}) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \quad (1)$$

onde π_k representa o peso da k -ésima componente, satisfazendo $\sum_{k=1}^K \pi_k = 1$, e $\boldsymbol{\mu}_k$ e $\boldsymbol{\Sigma}_k$ correspondem, respectivamente, ao vetor de médias e à matriz de covariância da componente k . Os parâmetros são encontrados por meio da maximização de expectativa, alternando entre momento de estimação das probabilidades e atualização dos parâmetros.

O GMM é um algoritmo de soft clustering, ou seja, atribui probabilidades ao invés de classificar unicamente a um dos componentes, tornando-o mais robusto. Serão classificadas como anomalias os dados que possuam baixa probabilidade (log-verossimilhança).

2) *Justificativa:* O GMM foi escolhido primeiramente por ser um modelo de clustering, uma das abordagens mais importantes para aprendizagem não-supervisionada. Entre os algoritmos desse tipo, o GMM se destaca por utilizar o soft clustering probabilístico e modelar os grupos com base em múltiplas distribuições, lidando bem com dados mais complexos.

Por fim, o GMM apresenta uma boa relação entre expressividade e custo computacional, sendo capaz de modelar correlações entre variáveis (por meio de matrizes de covariância completas) sem exigir grandes volumes de dados ou recursos computacionais elevados. Dessa forma, o GMM se configura como uma escolha sólida e complementar aos modelos de isolamento e aprendizado profundo considerados neste trabalho, contribuindo para uma comparação abrangente entre diferentes paradigmas de detecção de anomalias aplicados ao dataset em questão.

3) *Espaço de Busca*: Para a tunagem foi utilizado o grid search. Os parâmetros buscados para descobrir o melhor número de componentes foram limitados a números baixos devido à natureza do problema: o braço robótico, que deve seguir movimentos repetitivos em condições normais, não precisará de muitos clusters, como foi verificado pelo seu melhor parâmetro.

TABLE III
ESPAÇO DE BUSCA DOS HIPERPARÂMETROS UTILIZADOS NA TUNAGEM DO MODELO GAUSSIAN MIXTURE MODEL (GMM).

Categoria	Hiperparâmetro e valores considerados
2*Janelamento	(new_window_size): {30, 60} (window_overlap): {0, 15}
4*GMM	(n_components): {1, 2, 3, 4} (covariance_type): {diag, full} (reg_covar): {1e-5} (threshold_percentile): {1, 2.5, 5}

4) *Hiperparâmetros Seleccionados*: Para a janela, o melhor foi com tamanho 60 e overlap 0. Para o modelo, foi 1 componente, tipo de covariância 'diag' e limiar 1.

B. Isolation Forest

1) *Conceitos Básicos*: O Isolation Forest é um modelo de aprendizagem não-supervisionada bastante relacionado com detecção de anomalias. Diferentes de outros modelos que também são frequentemente usados para essa área, o Isolation Forest utiliza o princípio de isolamento, isto é, dado que anomalias devem ser bem diferenciáveis dos dados normais, elas serão mais fáceis de isolar com um particionamento aleatório.

O algoritmo constrói um conjunto de árvores binárias, denominadas isolation trees, nas quais cada nó interno realiza uma divisão aleatória escolhendo um atributo e um valor de corte também aleatório dentro do intervalo observado desse atributo. O processo continua recursivamente até que a instância seja isolada em um nó folha ou até que seja atingida uma profundidade máxima.

Para definir se um dado é anômalo ou não, o modelo utiliza o comprimento médio do caminho, ou seja, a profundidade média necessária para isolá-lo. O anomaly score gerado parte do princípio que dados anômalos serão isolados mais cedo que os demais.

2) *Justificativa*: Primeiramente, esse modelo foi projetado precisamente para problemas como o nosso: não-supervisionado, com detecção de anomalias raras, e utiliza uma abordagem diferente da dos outros modelos escolhidos.

Além disso, ele funciona como um bom contraponto para o clustering previamente testado: datasets como o que usamos, relacionado à sensores industriais, costumam ter muitas dimensões, especialmente com o janelamento temporal, o que impacta negativamente modelos como K-Means e GMM, a chamada "maldição da dimensionalidade". Por não depender de métricas de distância globais, o Isolation Forest evita esse problema. O caráter aleatório e de ensemble dele também auxilia na detecção de anomalias de vários tipos, como é o caso desse dataset.

3) *Espaço de Busca*: a tunagem foi feita utilizando o tradicional método de grid search, testando todas as combinações dos parâmetros possíveis detalhados na tabela.

TABLE IV
ESPAÇO DE BUSCA DOS HIPERPARÂMETROS UTILIZADOS NA TUNAGEM DO MODELO ISOLATION FOREST.

Categoria	Hiperparâmetro e valores considerados
2*Janelamento	(new_window_size): {30, 60} (window_overlap): {0, 15}
3*Isolation Forest	(n_estimators): {50, 100, 150, 200, 250, 300} (max_samples): {auto, 0.25, 0.5, 0.75} (contamination): {auto, 0.01, 0.05, 0.1}

4) *Hiperparâmetros Seleccionados*: para a janela, o tamanho e o overlap foram 60 e 15, respectivamente. No modelo, os melhores parâmetros foram 50 árvores, 0.75 fração de amostras por árvore e taxa de contaminação automática ('auto').

C. Autoencoder Convolutacional

1) *Conceitos Básicos*: Autoencoders são modelos de aprendizado não supervisionado baseados em redes neurais artificiais cujo objetivo é aprender uma representação compacta dos dados por meio da reconstrução da entrada. Formalmente, um autoencoder é composto por duas funções principais: um codificador, que projeta os dados de entrada em um espaço latente de menor dimensionalidade, e um decodificador, responsável por reconstruir os dados originais a partir dessa representação comprimida. O treinamento é realizado minimizando uma função de perda de reconstrução, tipicamente o erro quadrático médio (MSE), entre a entrada e sua reconstrução.

No caso de Autoencoders Convolutacionais 1D, camadas densas são substituídas por camadas convolutacionais unidimensionais, que exploram explicitamente a estrutura temporal ou sequencial dos dados. As convoluções 1D aplicam filtros locais ao longo do eixo temporal, permitindo ao modelo capturar padrões de curto alcance, como tendências locais, transições abruptas e correlações temporais entre variáveis sensoriais, auxiliado por operações como pooling e upsampling, permitindo reduzir e recuperar as representações.

Para detecção de anomalias, o autoencoder é treinado exclusivamente com dados considerados normais, aprendendo a reconstruir adequadamente padrões recorrentes do comportamento normal, enquanto apresenta maior erro de reconstrução para amostras divergentes. Assim, o erro de reconstrução passa a ser utilizado como uma métrica contínua de anomalia, permitindo a definição de um limiar para classificação.

2) *Justificativa*: A escolha de um Autoencoder Convolutacional 1D para este projeto se justifica principalmente pela natureza temporal e multivariada dos dados analisados. Diferentemente de métodos baseados apenas em distribuição estatística ou isolamento geométrico, autoencoders convolutacionais são capazes de aprender relações não lineares complexas entre múltiplos sensores ao longo do tempo, sem a necessidade de rótulos explícitos. Além disso, o uso de

convoluções 1D confere ao modelo invariância local e capacidade de generalização, algo muito valioso em ambientes que podem ter ruído.

É um modelo bastante robusto, baseado em aprendizagem profunda, capaz de modelar relações complexas e altamente não-lineares.

Por fim, a abordagem baseada em erro de reconstrução permite uma separação clara entre dados normais e anômalos, como evidenciado pelos resultados experimentais obtidos neste trabalho, nos quais as anomalias apresentaram erros de reconstrução significativamente superiores aos observados em dados normais. Essa característica torna o Autoencoder Convolutacional uma ferramenta eficaz e robusta para detecção de anomalias em séries temporais multivariadas.

3) *Espaço de Busca*: Foi feito um grid search para tunagem de hiperparâmetros. Antes da tunagem propriamente dita, diversas arquiteturas foram analisadas manualmente para delimitar nosso espaço de busca, evitando uma grid search muito extensa e demorada, testando configurações mais complexas e com mais camadas, chegando até a 64 filtros. No entanto, elas apresentaram resultados semelhantes a arquiteturas mais simples de cerca de 8 filtros. Para evitar overfitting e custo computacional elevado, optamos pelos parâmetros abaixo.

TABLE V
ESPAÇO DE BUSCA DOS HIPERPARÂMETROS UTILIZADOS NA TUNAGEM DO AUTOENCODER CONVOLUCIONAL 1D.

Categoria	Hiperparâmetro e valores considerados
5*Autoencoder Convolutacional 1D	(filters): {{8,8}, {8,6}, {12,8}} (kernel_sizes): {{3,3}, {5,3}} (latent_channels): {4, 8} (batch_size): {64} (epochs): {60} (lr): { 10^{-3} } (patience): {5} (percentile): {95.0, 97.5, 99}

4) Hiperparâmetros Selecionados:

IV. ANÁLISE E COMPARAÇÃO DE RESULTADOS

Deverá conter as métricas que foram utilizadas para a análise juntamente com revisão de conceito e justificativas. É fundamental comparar os resultados obtidos entre os diferentes modelos treinados. Além disso, é interessante utilizar ferramentas estatísticas e/ou testes de hipótese quando cabível.

TABLE VI
RESULTADOS DE DESEMPENHO DO MODELO GAUSSIAN MIXTURE MODEL (GMM) NO CONJUNTO DE TESTE.

Métrica	Valor
Acurácia (Accuracy)	0.8256
Precisão (Anomalia como positiva)	0.1899
Revocação (Recall – Anomalia)	1.0000
F1-score (Anomalia)	0.3192
ROC AUC	1.0000
PR AUC (Average Precision)	1.0000
Média de log-verossimilhança (normal)	-1672.9433
Média de log-verossimilhança (anômalo)	-12554031.9700

TABLE VII
MATRIZ DE CONFUSÃO DO MODELO GAUSSIAN MIXTURE MODEL (GMM) NO CONJUNTO DE TESTE.

2ªClasse Real	Classe Predita	
	Anomalia	Normal
Anomalia	166	0
Normal	708	3186

O Gaussian Mixture Model (GMM) alcançou recall igual a 1,0, detectando todas as anomalias presentes no conjunto de teste, o que evidencia elevada sensibilidade. No entanto, essa capacidade foi acompanhada por um número expressivo de falsos positivos, refletido em uma precisão de aproximadamente 0,19 e em um F1-score de cerca de 0,32, ainda que superior ao obtido pelo Isolation Forest. Apesar desse desempenho binário moderado, as métricas baseadas em ranqueamento revelam um cenário significativamente mais favorável: tanto a ROC AUC quanto a PR AUC atingiram o valor máximo (1,0). Essa discrepância indica que o GMM foi capaz de separar perfeitamente amostras normais e anômalas no espaço de escores contínuos (log-verossimilhança), mas que o limiar adotado para a decisão final não foi ótimo. Tal interpretação é corroborada pela diferença extrema observada entre as log-verossimilhanças médias das amostras normais e anômalas, sugerindo que as anomalias apresentam padrões estatísticos profundamente incompatíveis com a distribuição aprendida a partir dos dados normais.

A escolha do número de componentes igual a um para o melhor GMM reforça essa interpretação. Um modelo unimodal assume que os dados normais podem ser adequadamente descritos por uma única distribuição Gaussiana, tratando desvios significativos como anômalos. Nesse contexto, a excelente separação observada nos scores sugere que o conjunto de dados normal é relativamente homogêneo, enquanto as anomalias representam desvios extremos, o que favorece abordagens baseadas em densidade. Ainda assim, a taxa de falsos positivos indica a necessidade de uma estratégia de calibração de limiar mais criteriosa, por exemplo, baseada na maximização do F1-score ou em restrições de precisão mínima no conjunto de validação.

TABLE VIII
RESULTADOS DE DESEMPENHO DO MODELO ISOLATION FOREST NO CONJUNTO DE TESTE.

Métrica	Valor
Acurácia (Accuracy)	0.8579
Precisão (Anomalia como positiva)	0.1487
Recall – Anomalia	0.5241
F1-score (Anomalia)	0.2317
ROC AUC	0.7745
PR AUC (Average Precision)	0.3212

O Isolation Forest apresentou acurácia global de aproximadamente 85,8%, valor influenciado principalmente pelo forte desbalanceamento entre amostras normais e anômalas. Apesar dessa acurácia relativamente elevada, o modelo obteve baixa precisão para a classe anômala (cerca de 0,15), indicando

TABLE IX
MATRIZ DE CONFUSÃO DO MODELO ISOLATION FOREST NO CONJUNTO DE TESTE.

2*Classe Real	Classe Predita	
	Anomalia	Normal
Anomalia	87	79
Normal	498	3396

que a maioria das amostras sinalizadas como anômalas corresponde, na realidade, a padrões normais. Em contrapartida, o recall de aproximadamente 0,52 sugere que o modelo foi capaz de identificar pouco mais da metade das anomalias reais presentes no conjunto de teste. Esse comportamento resulta em um F1-score reduzido (cerca de 0,23), evidenciando um compromisso desfavorável entre sensibilidade e precisão. A métrica ROC AUC de 0,77 indica que, embora o modelo possua capacidade razoável de ranqueamento entre padrões normais e anômalos, sua fronteira de decisão não é suficientemente discriminativa para fornecer decisões binárias confiáveis sem uma calibração adicional do limiar de decisão. Somado à isso, o isolation forest foi o único modelo que classificou algumas anomalias como normais, algo que seria bastante problemático no ambiente de produção.

V. CONCLUSÃO E DISCUSSÃO

Explicar os principais achados ao longo do trabalho bem como vantagens e limitações de métodos e/ou algoritmos selecionados. Além disso, apresentar principais *insights* extraídos e potenciais trabalhos futuros.

REFERÊNCIAS

- [1] H. Kayan, R. Heartfield, O. Rana, P. Burnap, and C. Perera, “CASPER: Context-Aware IoT Anomaly Detection System for Industrial Robotic Arms,” *ACM Transactions on Internet of Things*, vol. 5, no. 3, art. 18, Aug. 2024, doi: 10.1145/3670414.
- [2] H. Kayan, R. Heartfield, O. Rana, P. Burnap, and C. Perera, “Real-Time Anomaly Detection for Industrial Robotic Arms Using Edge Computing,” *2023 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 2023.
- [3] A. Blázquez-García, M. S. Del Aguila, I. López, P. López, and F. Sánchez, “A review on outlier/anomaly detection in time series data,” *ACM Computing Surveys*, vol. 54, no. 3, pp. 1–33, 2021.
- [4] Y. Yang et al., “Unsupervised Anomaly Detection for Autonomous Robots via Mahalanobis SVDD with Audio-IMU Fusion,” *arXiv preprint arXiv:2505.05811*, 2025.
- [5] S. Chirayil Nandakumar et al., “Anomaly detection methods in autonomous robotic missions,” *Sensors*, vol. 24, no. 4, p. 1330, 2024.
- [6] Keogh, E., Lin, J. Clustering of time-series subsequences is meaningless: implications for previous and future research. *Knowl Inf Syst* 8, 154–177 (2005).
- [7] Keogh, E., Lin, J. Clustering of time-series subsequences is meaningless: implications for previous and future research. *Knowl Inf Syst* 8, 154–177 (2005).