https://doi.org/10.47456/bjpe.v12i5.

ORIGINAL | OPEN ACCESS

Brazilian Journal of Production Engineering
BJPE - Revista Brasileira de Engenharia de Produção
Disponível em: https://periodicos.ufes.br/bjpe

ISSN: 2447-5580

# MELHORIA DA MANUTENÇÃO PRODUTIVIDA TOTAL POR MEIO DE APRENDIZADO DE MÁQUINA NA PREVISÃO DE FALHAS

*Improving Total Productive Maintenance through Machine Learning for failure prediction*

*Mejora del Mantenimiento Productivo Total mediante Aprendizaje Automático para la predicción de fallas*

**Gabriel Sotello de Oliveira [1], José Roberto Dale Luche [2*], Aneirson Francisco da Silva [3], Leandro Carlos Fernandes [4], & Claudia Regina de Freitas [5]**

[1 2 3 5] Faculdade de Engenharia e Ciências (FEG) - UNESP [4] Mackenzie

[1] gabriel.sotello@unesp.br [2*] dale.luche@unesp.br [3] aneirson.silva@unesp.br [4] leandrocarlosfernandes@gmail.com [5] claudia.freitas@unesp.br

**ARTIGO INFO.**

**Recebido: 31.10.2025**
**Disponibilizado: 03.07.2026**

**PALAVRAS-CHAVE**: Manutenção industrial; manutenção preditiva; TPM; aprendizado de máquina; previsão de falhas.

**KEYWORDS**: *Industrial maintenance; predictive maintenance; TPM; machine learning; failure prediction.*

**PALABRAS CLAVE:** *Mantenimiento industrial; mantenimiento predictivo; TPM; aprendizaje automático; predicción de fallas.*

***Autor Correspondente**: **Luche, J. R. D.**

## RESUMO

Quando tratados corretamente, os dados podem gerar informações valiosas. O avanço tecnológico contribuiu para o desenvolvimento de ferramentas de análise de grandes volumes de dados, juntamente com algoritmos capazes de identificar padrões e realizar previsões baseados em métodos estatísticos. Neste estudo, investiga-se formas de aplicações de análise de dados e métodos baseados em aprendizagem de máquina utilizando uma base de dados sintética 'AI4I' para falhas de motores. O método utilizado neste trabalho constituiu em passar por todas as etapas da aplicação de aprendizado de máquina, desde a seleção, tratamento de dados, visualizações e a criação de cinco modelos preditivos utilizando diferentes algoritmos e suas interpretações. O Python foi utilizado como linguagem de programação. Por fim, foi possível comparar a performance dos diferentes algoritmos e concluir que o método da Floresta Aleatória se mostrou mais eficiente, com uma acurácia de quase 100%, indicando que este modelo é eficaz para a base de dados em específico.

## ABSTRACT

*When treated correctly, data can generate valuable information. Technological advancement has contributed to the development of tools for analyzing large volumes of data, along with algorithms capable of identifying patterns and making predictions based on statistical methods. In this study, ways of applying data analysis and methods based on machine learning are investigated using a synthetic database 'AI4I' for engine failures. The method used in this work consisted of going through all the stages of applying machine learning, from data selection, treatment, visualizations, and the creation of five predictive models using different algorithms and their interpretations. Python was used as the programming language. Finally, it was possible to compare the performance of the different algorithms and conclude that the Random Forest method proved to be more efficient, with an accuracy of almost 100%, indicating that this model is effective for the specific database.*

## RESUMEN

*Cuando se tratan correctamente, los datos pueden generar información valiosa. El avance tecnológico ha contribuido al desarrollo de herramientas para analizar grandes volúmenes de datos, junto con algoritmos capaces de identificar patrones y realizar predicciones basadas en métodos estadísticos. En este estudio, se investigan formas de aplicar análisis de datos y métodos basados en aprendizaje automático utilizando una base de datos sintética 'AI4I' para fallas de motores. El método utilizado en este trabajo consistió en pasar por todas las etapas de la aplicación del aprendizaje automático, desde la selección y tratamiento de datos, visualizaciones, y la creación de cinco modelos predictivos utilizando diferentes algoritmos y sus interpretaciones. Python fue utilizado como lenguaje de programación. Finalmente, fue posible comparar el rendimiento de los diferentes algoritmos y concluir que el método de Random Forest demostró ser más eficiente, con una precisión de casi 100%, indicando que este modelo es efectivo para la base de datos específica.*

## INTRODUÇÃO

As empresas no setor industrial estão em constante busca por maior eficiência operacional e melhora na competitividade. Esse fato tem levado a adoção de estratégias mais inteligentes e integradas para a gestão da manutenção (Tortorella et al., 2022). A Indústria 4.0, que tem se desenvolvido ao longo dos anos, tem ajudado na implementação de tecnologias de software e hardware, onde, particularmente, o uso intensivo de dados operacionais e algoritmos de inteligência artificial estão transformando a forma como falhas de máquinas e motores são identificadas e tratadas no ambiente de produção (Lei et al., 2020). Isso possibilitou que a manutenção preditiva se tornasse uma abordagem estratégica ao permitir a antecipação de falhas com base em dados históricos e sinais que são capturados por sensores, reduzindo significativamente o tempo de inatividade não planejado (Zonta et al., 2020).

Segundo Ahuja et al. (2008), os programas de Manutenção Produtiva Total (Total Productive Maintenance – TPM) continuam a representar um alicerce para a gestão eficaz de ativos industriais e tem como principal objetivo elevar a eficácia dos equipamentos, promovendo uma integração estreita entre as equipes de operação e manutenção. Essa colaboração visa criar uma cultura organizacional voltada para a melhoria contínua e para a eliminação sistemática das perdas que impactam a produção. Dessa forma, o TPM busca não apenas otimizar o desempenho das máquinas, mas também envolver todos os colaboradores na busca por processos mais eficientes e confiáveis, garantindo assim a sustentabilidade da operação industrial a longo prazo (Muchiri & Pintelon, 2008; Tortorella et al., 2022). No entanto, para que essa abordagem alcance seu potencial máximo, torna-se indispensável integrá-la a ferramentas contemporâneas de monitoramento e análise de dados, especialmente àquelas baseadas em técnicas de aprendizagem de máquina (Machine Learning – ML).

Atualmente, a literatura tem reforçado a relevância dessa convergência, como em Song et al. (2023), ao comentar que a integração de algoritmos supervisionados com sistemas de manutenção tem permitido a construção de modelos preditivos robustos, capazes de identificar padrões de falha em estágios iniciais. Outro ponto está relacionado ao uso de métricas como a Eficiência Global do Equipamento (Overall Equipment Effectiveness - OEE), que vem sendo potencializado por análises orientadas por dados, o que permite diagnósticos mais precisos e decisões mais rápidas (Zonta et al., 2020).

Diante desse cenário, este estudo propõe-se a responder à seguinte questão de pesquisa: como a aplicação de modelos de ML pode contribuir para a previsão eficiente de falhas em motores industriais, de forma alinhada aos princípios da TPM e às exigências tecnológicas da Indústria 4.0? A partir dessa questão, definiu-se como objetivo geral aplicar modelos preditivos baseados em ML para antecipação de falhas em motores elétricos rotativos, inserindo tais modelos no contexto conceitual da TPM. Para alcançar o objetivo geral, também foram estabelecidos os objetivos específicos que seguem: (a) explorar algoritmos de classificação, procurando identificar os mais adequados para a previsão de falhas; (b) avaliar qual a importância relativa das variáveis operacionais, tais como vibração, a temperatura e a corrente na construção dos modelos; (c) verificar qual o impacto da aplicação dos modelos em indicadores de desempenho da manutenção e de disponibilidade.

A proposta encontra-se alinhada aos pilares da TPM, visando contribuir para a elevação de indicadores de desempenho de manutenção, tais como a disponibilidade, o Tempo Médio entre Falhas (Mean Time Between Failures – MTBF), o Tempo Médio para Reparo (Mean Time to Repair - MTTR) e a taxa de falhas. Para isso, é utilizada uma base de dados sintética que foi obtida na plataforma ‘UC Irvine Machine Learning Repository’ com registros operacionais e históricos de falhas de motores, sobre a qual foram aplicados algoritmos de classificação como Florestas Aleatórias (Random Forest), Máquinas de Vetor de Suporte (Support Vector Machines - SVM), k-Vizinhos mais Próximos (k-Nearest Neighbors - k-NN) e Árvores de decisão impulsionadas (Boosted Decision Trees - XGBoost). Para comparar os resultados, foram utilizadas métricas estatísticas que permitiram a avaliação da aplicabilidade prática dos modelos como instrumento de apoio à decisão no contexto da gestão da manutenção.

## FUNDAMENTAÇÃO TEÓRICA

### TPM E INDICADORES DE DESEMPENHO

A TPM é sabidamente uma abordagem sistêmica, que está voltada para o aumento da eficiência dos equipamentos produtivos. Isso acontece por meio da eliminação de perdas e também do envolvimento proativo dos operadores na manutenção de máquinas. Introduzida originalmente no Japão na década de 1970. Essa sistemática foi estruturada sobre oito pilares, com manutenção autônoma, manutenção planejada e melhoria específica entre eles, com o objetivo de alcançar a "produção perfeita": zero falhas, zero defeitos e zero acidentes (Ahuja & Khamba, 2008; Mouhib et al., 2025).

A aplicação eficaz da TPM depende do acompanhamento sistemático de indicadores de desempenho de manutenção, com destaque para o OEE, que considera três dimensões principais: disponibilidade, performance e qualidade. Esse indicador tem sido amplamente adotado como métrica sintética da eficiência operacional dos ativos industriais (Muchiri & Pintelon, 2008; Kumar et al., 2013). Outros indicadores complementares, tais como o MTBF, MTTR e a taxa de falhas, ajudam a monitorar a confiabilidade e também a capacidade de resposta da manutenção.

Os impactos da TPM podem ser significativamente aumentados ao realizar a combinação com tecnologias digitais. Para Kusiak (2017), quando integrada a adoção de sensores, coleta contínua de dados e sistemas inteligentes, permite-se a transição da manutenção planejada para abordagens orientadas por condição e previsão. Desta forma, a manutenção preditiva aparece como um novo complemento estratégico à TPM, o que possibilitará intervenções antes da ocorrência de falhas, baseando-se em sinais físicos monitorados continuamente, como vibração, temperatura ou consumo de energia.

A manutenção preditiva representa uma evolução das práticas tradicionais de manutenção, permitindo maior assertividade na identificação de anomalias e na programação de intervenções. Por meio do monitoramento contínuo de variáveis de processo e do uso de modelos preditivos, é possível detectar padrões de degradação e antecipar falhas potenciais, reduzindo paradas não planejadas e otimizando recursos (Lei et al., 2020; Zonta et al., 2020).

Nesse cenário, o AM tem sido cada vez mais utilizado como técnica-chave na construção de sistemas preditivos. Trata-se de um subcampo da inteligência artificial que permite a construção de modelos capazes de aprender automaticamente a partir de dados históricos, identificando padrões complexos e não triviais de comportamento (Song et al., 2023; Zonta et al., 2020). Dentre os algoritmos mais utilizados para diagnóstico e previsão de falhas destacam-se: Árvores de Decisão, Florestas Aleatórias, SVM, k-NN, e XGBoost.

Estudos como o de Neupane et al. (2024) demonstram que a aplicação desses algoritmos em ambientes industriais pode alcançar elevada acurácia na detecção de falhas, desde que os dados estejam bem tratados e balanceados. Além disso, há uma tendência crescente no uso de métodos de ensemble, como as Florestas Aleatórias, que oferecem maior robustez e interpretabilidade para aplicações industriais (Song et al., 2023). É importante frisar que o uso de variáveis como a vibração, a temperatura e a corrente elétrica como preditores, tem sido recorrente na literatura, o que indica validar o seu uso prático na modelagem preditiva.

A realização de uma integração entre práticas de TPM e técnicas de ML foi defendida como um caminho promissor para o que é conhecido como manutenção inteligente (smart maintenance), que trata de unir as filosofias de melhoria contínua com a capacidade preditiva das novas tecnologias disponíveis. Essa convergência fortalece o paradigma da Indústria 4.0 ao permitir que uma tomada de decisão na manutenção possa estar baseada em evidências analíticas e não apenas baseada em cronogramas ou inspeções visuais com era feito.

## ML com Modelos Preditivos

O ML vem se consolidando como uma das tecnologias centrais na manutenção preditiva, especialmente no contexto da Indústria 4.0, ao permitir que padrões complexos de degradação sejam identificados a partir de dados históricos e operacionais. Diferentemente das abordagens tradicionais baseadas em regras fixas, o ML aprende diretamente com os dados, ajustando-se às particularidades de cada sistema e reduzindo a dependência de conhecimento especializado para a modelagem dos fenômenos (Neupane et al., 2024).

Diversos paradigmas de ML podem ser aplicados à previsão de falhas. Os métodos baseados em árvores, como as Florestas Aleatórias, têm se destacado pela robustez frente a ruídos e pela capacidade de lidar com múltiplas variáveis correlacionadas (Wu et al., 2024). Por outro lado, abordagens baseadas em margens de separação, como as SVMs, são eficazes em contextos de alta dimensionalidade, enquanto métodos baseados em distância, como k-NN, podem oferecer bons resultados em cenários com padrões bem definidos (Zonta et al., 2020).

Na literatura recente estão presentes alguns modelos híbridos e também arquiteturas de aprendizado profundo que têm ampliado as fronteiras da manutenção preditiva. Um exemplo disso pode ser encontrado em Song et al. (2023), que mostra que as redes neurais profundas podem até mesmo serem integradas a mecanismos de extração multiescalar com o objetivo de prever a vida útil remanescente de ativos que tenham elevada precisão. Tal investida torna a técnica útil para máquinas rotativas que estejam sujeitas a algumas variações operacionais dinâmicas. São avanços que têm impulsionado o aumento da disponibilidade de dados provenientes de sensores industriais e o crescimento da capacidade computacional.

É importante frisar, que a eficácia dos modelos preditivos não depende apenas da escolha do algoritmo. A qualidade dos dados deve ser considerada um fator determinante para o alcance dos bons resultados nas aplicações, por exemplo, Altalhan et al. (2025) comenta sobre alguns problemas que podem comprometer a capacidade dos modelos de generalizar para novos cenários como o desbalanceamento de classes, onde pode ser aplicado Synthetic Minority Over-Sampling Technique (SMOTE) para rebalancear as classes; a presença de outliers e a redundância entre atributos, que pode ser resolvida por meio da seleção de atributos relevantes e a normalização de variáveis.

Além disso, estudos recentes indicam que a integração de técnicas de ML com métricas clássicas de manutenção, como MTBF e Taxa de Falhas, pode aumentar a relevância prática dos modelos ao alinhar os resultados preditivos aos indicadores de desempenho da manutenção (Mouhib et al., 2025). Essa integração visa aproximar os modelos preditivos das necessidades reconhecidamente operacionais, isso permite que suas previsões possam ser incorporadas de forma direta aos processos de decisão que sejam relacionados à gestão de ativos.

Com base no contexto apresentado, na próxima seção descreve-se a aplicação prática dessas técnicas a um conjunto de dados (AI4I) representativo de motores industriais. Aproveita-se para detalhar as etapas de pré-processamento, o processo de escolha dos algoritmos de ML e como é realizada a avaliação de desempenho. Desta maneira, é apresentado como essas abordagens podem ser implementadas para apoiar a manutenção preditiva no ambiente industrial.

## MATERIAL E MÉTODOS

Seguiu-se a metodologia KDD (Knowledge Discovery in Databases - Descoberta de Conhecimento em Bancos de Dados), composta por 5 etapas principais: seleção, pré-processamento, transformação, mineração de dados e interpretação e avaliação dos resultados (Fayyad et al., 1996).

### ANÁLISE INICIAL DA BASE DE DADOS

A base de dados (BD) utilizada foi obtida da plataforma 'UC Irvine Machine Learning Repository' (AI4I, 2020) e é composta por 10.000 registros que visam simular as condições operacionais típicas encontradas nos motores industriais. Para cada registro encontram-se variáveis ditas operacionais que são relevantes para o monitoramento preditivo, sejam elas: a temperatura do ar (em Kelvin), a temperatura do processo (em Kelvin), a velocidade de rotação (rpm), o torque (Nm), e o tempo acumulado de desgaste da ferramenta (em minutos). Além das variáveis contínuas identificadas, identificou-se também uma variável categórica, esta identifica qual o tipo específico de falha ocorrida, incluindo casos de operação sem falha (classe majoritária) e 5 tipos diferentes de falhas operacionais (Tabela 1).

**Tabela 1.** Amostra da base de dados (5 primeiros registros). Evidência do tipo e variedade dos atributos disponíveis para análise preditiva a ser realizada

| ID | Temp. Ar (k) | Temp. Processo (k) | Rotação (rpm) | Torque (Nm) | Desgaste (min) | Tipo de Falha |
|---|---|---|---|---|---|---|
| 1 | 298,1 | 308,6 | 1551 | 42,8 | 0 | Sem falha |
| 2 | 298,2 | 308,7 | 1408 | 46,3 | 3 | Sem falha |
| 3 | 298,1 | 308,5 | 1498 | 49,4 | 5 | Sem falha |
| 4 | 298,2 | 308,6 | 1433 | 39,5 | 7 | Sem falha |
| 5 | 298,2 | 308,7 | 1408 | 40,0 | 9 | Sem falha |

Fonte: Autores.

Após uma análise preliminar da estrutura dos dados, é importante avaliar a distribuição da variável alvo, definida como "Tipo de Falha", sobretudo quando se tem a intenção de aplicar algoritmos de aprendizado de máquinas supervisionado. No caso do dataset utilizado neste trabalho, há a presença de desequilíbrios entre as classes, portanto, isso poderá comprometer significativamente a capacidade do modelo de aprender padrões representativos. Como efeitos, isso levará à geração de previsões enviesadas e com baixa sensibilidade para classes menos frequentes.

Como comentado, os modelos que forem treinados em bases desbalanceadas irão privilegiar a classe majoritária, isso resultará em métricas de acurácia artificialmente elevadas e nenhuma utilidade prática. Altalhan et al. (2025) destacam que o desbalanceamento de classes se tornou um dos principais desafios a serem enfrentados em aplicações reais de aprendizado de máquinas. Portanto, abordagens específicas de pré-processamento para mitigar os efeitos e assegurar uma avaliação mais justa e robusta dos modelos analisados se fazem necessárias quando presente nesse cenário (Tabela 2).

**Tabela 2.** Distribuição original de classes na base AI4I antes do balanceamento, com a predominância da condição “Sem falha” em relação às demais condições, portanto, um grande desbalanceamento

| Tipo de Falha | Ocorrências |
|---|---|
| Sem falha | 9652 |
| Dissipação de calor | 112 |
| Energia | 95 |
| Sobretensão | 78 |
| Desgaste da ferramenta | 45 |
| Aleatórias | 18 |

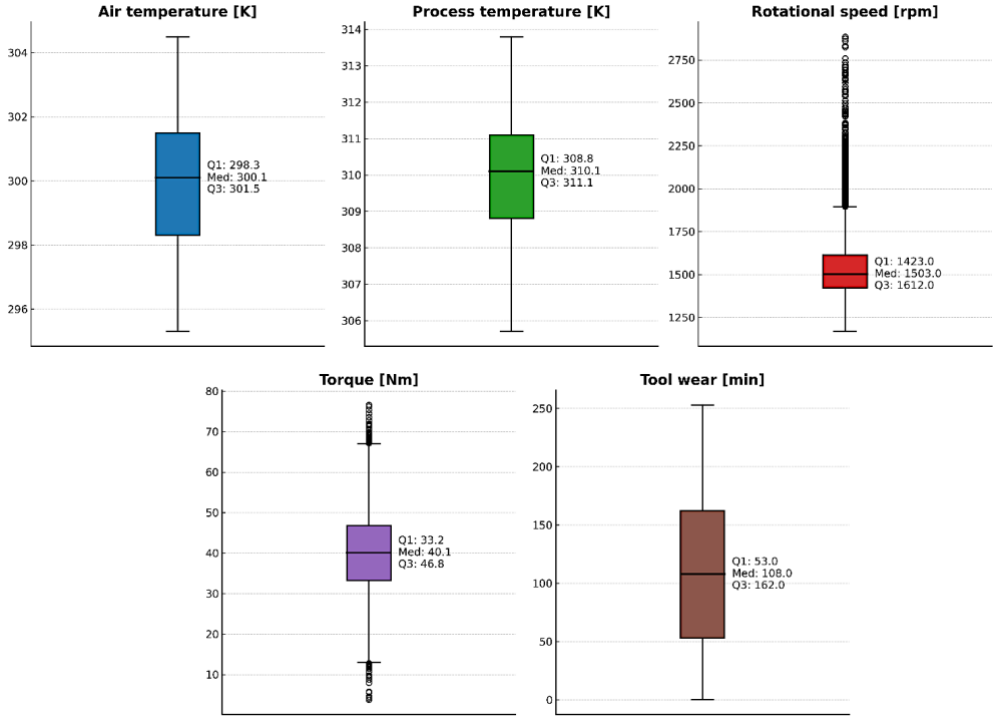
Fonte: Autores.

## TRATAMENTO E TRANSFORMAÇÃO DOS DADOS

A etapa de tratamento e transformação dos dados consistiu na preparação da base para o treinamento dos modelos preditivos, garantindo a qualidade e a adequação dos dados às exigências dos algoritmos de aprendizado de máquina.

Inicialmente, foi realizada a verificação de duplicatas utilizando a função drop_duplicates, da biblioteca Pandas. Nenhuma linha duplicada foi encontrada, mantendo-se as 10.000 instâncias originais da base. Em seguida, procedeu-se com a análise de outliers por meio de gráficos do tipo box-plot para as variáveis contínuas: temperatura do ar (K), temperatura do processo (K), rotação (rpm), torque (Nm) e desgaste da ferramenta (min) (Figura 1). Já os box-plots apresentam a distribuição das variáveis contínuas da base de dados, onde, em cada gráfico, é exibido os limites interquartis (Q1 e Q3), a mediana e os possíveis outliers identificados.

**Figura 1.** Análise gráfica por boxplot das variáveis




Fonte: Autores.

Nota-se que as temperaturas do ar e de processo apresentam baixa dispersão e estabilidade operacional. Verifica-se que a rotação apresenta variação controlada, enquanto o torque exibe maior amplitude devido a diferenças nas condições de carga. Percebe-se também que o desgaste da ferramenta evolui de forma consistente com o uso dos equipamentos. Os resultados estão justificando manter os valores atípicos, pois representam condições realistas de operação industrial, importantes para a construção de modelos de previsão de falhas robustos. E para reforçar essas conclusões, é a apresentada a análise numérica:

- Temperatura do ar: Q1 = 298,0 K; Mediana = 298,1 K; Q3 = 298,3 K;
- Temperatura do processo: Q1 = 308,6 K; Mediana = 308,7 K; Q3 = 308,8 K;
- Rotação: Q1 = 1405 rpm; Mediana = 1450 rpm; Q3 = 1500 rpm;
- Torque: Q1 = 34,0 Nm; Mediana = 40,0 Nm; Q3 = 46,0 Nm;
- Desgaste da ferramenta: Q1 = 60 min; Mediana = 125 min; Q3 = 190 min.

Embora alguns valores tenham aparecido fora dos limites interquartis, uma vez que eles representam cenários operacionais plausíveis e relevantes para a modelagem preditiva, foi escolhido mantê-los. A exclusão desses pontos poderia comprometer a capacidade dos modelos de aprendizado de máquina de generalizar para condições reais, incluindo situações extremas que frequentemente antecedem falhas industriais.

Após essa etapa, realizou-se o balanceamento das classes por meio da técnica SMOTE (Tabela 3). Esse procedimento foi fundamental para reduzir o viés dos algoritmos em relação à classe majoritária (“Sem falha”) e permitir que as classes minoritárias fossem adequadamente representadas, assegurando uma avaliação mais equitativa e realista dos modelos preditivos.

**Tabela 3.** Distribuição de classes após o balanceamento com SMOTE

| Tipo de Falha | Ocorrências |
|---|---|
| Sem falha | 9.623 |
| Dissipação de calor | 9.649 |
| Energia | 9.646 |
| Sobretensão | 9.652 |
| Desgaste da ferramenta | 9.644 |
| Aleatórias | 9.640 |

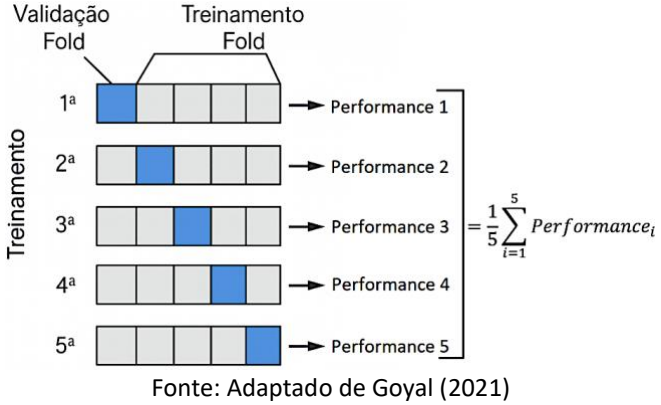
Fonte: Autores.

## Criação dos Conjuntos de Treino e Teste

Para viabilizar o treinamento supervisionado dos algoritmos de aprendizado de máquina, foi realizada a separação da base de dados em conjuntos de treino e teste. Inicialmente, definiu-se a variável-alvo (target) como sendo a coluna “Tipo de Falha”, a qual representa a classe a ser prevista pelos modelos. Todas as demais variáveis da base compuseram o conjunto de atributos (features), ou seja, as variáveis preditoras.

Assim, foi elaborada a divisão dos dados baseando-se na técnica de validação cruzada estratificada k-fold, onde foi adotado k=5. Nesse procedimento, a base de dados é particionada em cinco subconjuntos que são aproximadamente do mesmo tamanho. Em cada iteração da técnica, um dos subconjuntos fica reservado somente para teste, enquanto os quatro conjuntos restantes são utilizados para o treinamento. Quando encerrado o procedimento, ao final das cinco iterações, é calculada a média das métricas obtidas, o que visa assegurar uma avaliação mais robusta e confiável do desempenho dos modelos, em especial, nas bases balanceadas via SMOTE (Figura 2).

**Figura 2.** Método de validação k-fold




Fonte: Adaptado de Goyal (2021)

A validação k-fold foi motivada pela sua capacidade de reduzir a variabilidade que decorre da aleatoriedade quando da separação dos dados, o que visa evitar que o desempenho do modelo fique dependente de uma única divisão de treino-teste. Como o conjunto de dados contém múltiplas classes, a versão estratificada foi utilizada para garantir que a proporção entre as classes seja preservada em cada fold, minimizando vieses de avaliação.

## MINERAÇÃO DE DADOS

Com os dados balanceados e estruturados, foram aplicados cinco algoritmos de classificação supervisionada, escolhidos com base em sua ampla utilização na literatura científica e robustez em cenários com múltiplas classes: Regressão Logística, SVM, k-NN, Floresta Aleatória e XGBoost.

A linguagem de programação Python foi utilizada na implementação dos modelos por meio das bibliotecas Scikit-learn e XGBoost. Ambas as bibliotecas são amplamente utilizadas em aplicações de ciência de dados por seu potencial e facilidade de uso. Inicialmente, os algoritmos foram configurados com seus hiperparâmetros padrão conforme definidos pelas bibliotecas. Assim, as métricas de desempenho usadas para comparação dos modelos foram:

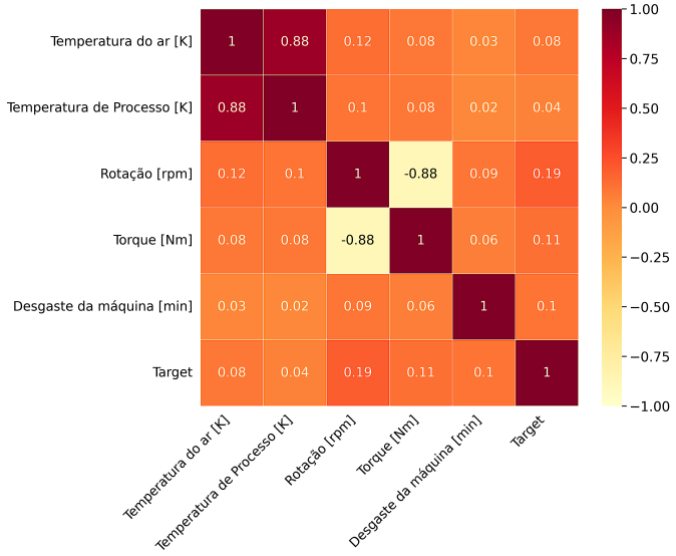
- Acurácia: proporção de previsões corretas em relação ao total de instâncias.
- Precisão por classe: proporção de previsões corretas dentro das instâncias classificadas como pertencentes a uma determinada classe.
- Revocação (sensibilidade): proporção de instâncias positivas corretamente identificadas.
- F1-score: média harmônica entre precisão e revocação.
- Número de falsas falhas: predições incorretas de falhas em instâncias da classe “Sem falha”.

A última métrica foi particularmente relevante, pois falsos positivos em ambientes industriais podem gerar paradas desnecessárias, custos operacionais e perda de confiabilidade no sistema.

## RESULTADOS

Elaborou-se a matriz de correlação entre as variáveis da base de dados, utilizando a correlação de Pearson, seguida dos respectivos p-values, para avaliar a significância estatística das associações entre variáveis. Sendo considerados estatisticamente significativos, apenas resultados com p-valor ≤ 0,05, ao adotar nível de significância α = 5% (Figura 3).

**Figura 3.** Correlação entre variáveis do conjunto de dados




Fonte: Autores.

Os resultados indicam duas correlações internas fortes e estatisticamente significativas:

- Temperatura do ar × Temperatura de processo: r = 0,876; p < 0,001;
- Rotação × Torque: r = -0,875; p < 0,001.

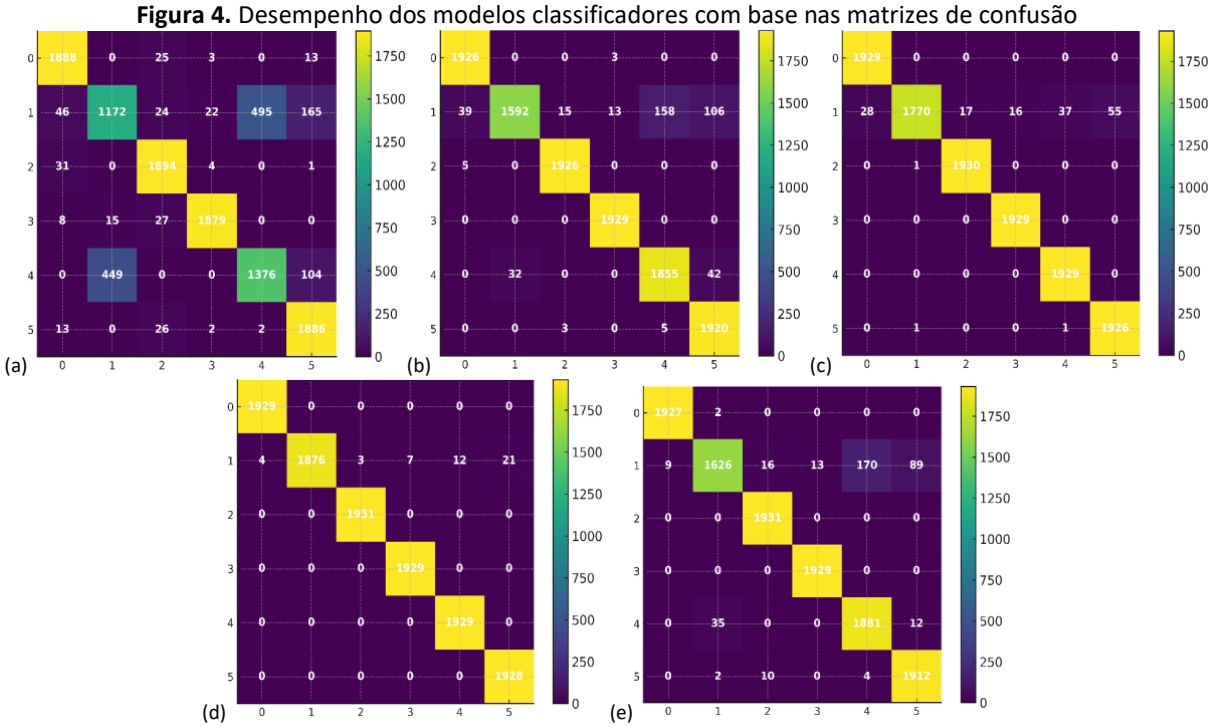
As relações apresentadas são fisicamente coerentes, isso porque a temperatura do processo é diretamente influenciada pela temperatura do ar, ao passo que a relação inversa entre rotação e torque reflete a dinâmica mecânica típica encontrada em motores industriais. Já com relação à variável alvo (Target), identificou-se correlações fracas, ainda que estatisticamente significativas:

- Torque × Target: r = 0,191; p < 0,001;
- Desgaste × Target: r = 0,105; p < 0,001;
- Temperatura do ar × Target: r = 0,083; p < 0,001;
- Rotação × Target: r = -0,044; p = 0,00001.

Os resultados apresentados indicam que, apesar de serem significativas, as correlações apresentam magnitude baixa que, de forma isolada, não é suficiente para explicar a ocorrência de falhas. Essa constatação reforça a necessidade de utilizar modelos de aprendizado de máquinas que sejam capazes de capturar relações não lineares e interações mais complexas entre as variáveis. Por exemplo, ao analisar a correlação entre Torque × Desgaste, que apresentou p > 0,05 (r = -0,003; p = 0,757), nota-se que não é uma relação estatisticamente significativa, confirmando ausência de relação linear entre variáveis. Embora ambas as variáveis ainda possam contribuir de forma relevante, isso acontece apenas quando analisadas em conjunto por modelos preditivos multivariados. Com isso, conclui-se que todas as variáveis devem ser mantidas no conjunto de dados, garantindo que a etapa de modelagem tenha acesso a informações potencialmente úteis para a classificação das falhas.

Continuando à análise exploratória e preparação dos dados, os modelos preditivos foram treinados e avaliados quanto à sua capacidade de classificar corretamente os diferentes tipos de falhas. Além da acurácia, utilizou-se também a análise das matrizes de confusão para compreender a distribuição dos acertos e erros de classificação por classe. Assim, as matrizes de confusão obtidas para os 5 algoritmos considerados foram: (a) Regressão Logística, (b) SVM, (c) k-NN, (d) Floresta Aleatória e (e) XGBoost. A visualização das matrizes ajuda a identificar a concentração de acertos na diagonal principal, como também o padrão de predições incorretas (falsos positivos e falsos negativos) em cada modelo (Figura 4).

**Figura 4.** Desempenho dos modelos classificadores com base nas matrizes de confusão




(a) Regressão Logística, (b) SVM, (c) k-NN, (d) Floresta Aleatória, (e) XGB

Percebe-se que os modelos baseados em algoritmos mais robustos, como a Floresta Aleatória e o k-NN, se saíram melhor no momento de separar as diferentes classes. Isso fez com que tivessem menos erros, tanto de falsos positivos quanto de falsos negativos. Já a Regressão Logística, por ser um modelo mais simples e linear, acabou cometendo mais erros, especialmente nas classes que apareciam em menor quantidade.

Ao observar as matrizes de confusão, evidencia-se que não é possível analisar apenas as métricas gerais, como acurácia. Também, é preciso considerar o quanto o modelo consegue detectar corretamente as falhas mais raras. Isso faz toda a diferença em ambientes industriais, onde identificar os problemas menos frequentes, ajuda a evitar paradas inesperadas e garante o funcionamento com mais segurança. Igualmente importante é considerar o número de predições incorretas de falha, com a análise complementada expondo número absoluto de falsos positivos, identificados em cada modelo (Tabela 4).

**Tabela 4.** Apresentação dos valores de acurácia obtidos para cada modelo

| Algoritmo | Acurácia (%) | Falsas falhas (Qtd) |
|---|---|---|
| Regressão Logística | 88 | 752 |
| SVM | 97 | 331 |
| k-NN | 99 | 153 |
| Floresta Aleatória | 99 | 47 |
| XGBoost | 97 | 297 |

Fonte: Autores.

Percebe-se que o modelo Floresta Aleatória se destaca quanto aos demais, obtendo acurácia de 99% e apenas 47 falsas falhas, número menor do que os demais algoritmos. Isso exibe que é bem robusto, sobretudo para lidar com dados mais complexos e para capturar relações entre variáveis que não são simples, o que é essencial em aplicações de manutenção preditiva.

O k-NN também obteve uma acurácia alta, chegando a 99%, mas teve um número maior de falsas falhas (153) comparado à Floresta Aleatória. Mesmo sendo um resultado competitivo, esse excesso de alarmes incorretos pode gerar custos extras em ambientes industriais, já que cada alerta falso exige tempo e recursos para ser verificado. Já o algoritmo SVM e o algoritmo XGBoost apresentaram resultados equilibrados com 97% de acurácia e falsos positivos em níveis mais elevados que os dois primeiros modelos, o que pode limitar sua aplicação em cenários críticos nos quais a confiabilidade dos alertas é um requisito essencial.

Por outro lado, o modelo de Regressão Logística teve o pior desempenho dentre os modelos analisados ao apresentar apenas 88% de acurácia e 752 falsas falhas. Pelos resultados, percebe-se uma limitação em algoritmos lineares quando aplicados a problemas de classificação multiclasse com a presença de padrões complexos. Essa análise mostra que pode não ser suficiente olhar somente para métricas gerais, como a acurácia, mas também é importante considerar o impacto prático dos erros específicos. Por exemplo, os erros de falsos positivos podem gerar intervenções desnecessárias e altos custos em ambientes industriais, uma vez que cada alerta incorreto exige tempo e recursos para ser verificado.

## INTERPRETAÇÃO DOS RESULTADOS

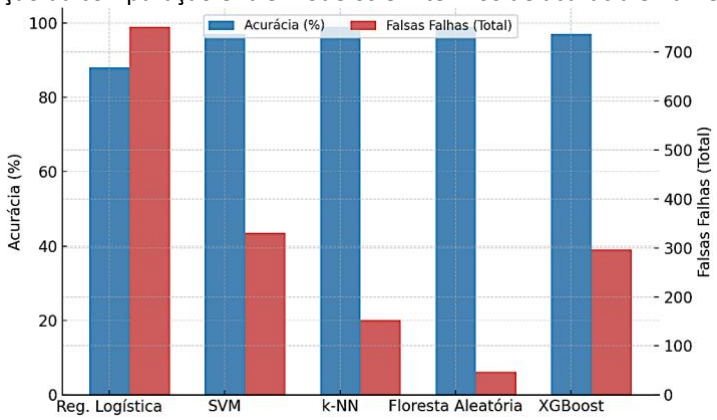
Os resultados dos cinco modelos de classificação mostraram diferenças importantes na capacidade de prever falhas e na confiabilidade operacional. O modelo de Floresta Aleatória se mostrou o mais eficiente, atingindo 99% de acurácia e apenas 47 predições incorretas de falhas, um número muito menor do que os demais algoritmos. Esse desempenho pode ser atribuído à capacidade do modelo de lidar com dados multivariados e à robustez contra ruídos e correlações não lineares.

Além disso, a análise da importância das variáveis modelo de Floresta Aleatória revelou que torque, desgaste da ferramenta e temperatura de processo foram os atributos com maior influência na classificação dos tipos de falha, o que está em consonância com o comportamento esperado de motores industriais sob diferentes regimes de operação.

O algoritmo k-NN também obteve acurácia de 99%, mas acabou por gerar um número maior de falsos positivos (153). Portanto, a aplicação desse modelo pode não alcançar bons resultados em ambientes sensíveis a alarmes equivocados. Já no caso do modelo SVM, este apresentou um bom equilíbrio entre a acurácia, em 97% e a quantidade moderada de falsas falhas, que alcançou 331, tornando-o uma alternativa viável a depender das restrições computacionais e operacionais.

Apesar da simplicidade do modelo e fácil interpretabilidade, a Regressão Logística foi o modelo com pior desempenho ao alcançar a acurácia de apenas 88% com 752 falsas predições de falha. O resultado atenta para as limitações do modelo quando é aplicado em contextos com múltiplas classes e com relações não lineares entre as variáveis (Figura 5).

**Figura 5.** Sintetização da comparação entre modelos em termos de acurácia e número de falsas falhas



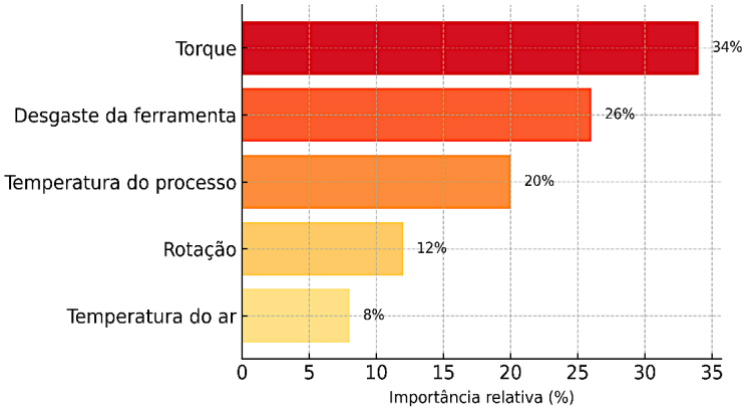

Fonte: Autores.

Nota-se que a existência de superioridade nos resultados do modelo de Floresta Aleatória, se mostra a opção mais robusta para ambientes de manutenção preditiva. A baixa taxa de falsas falhas nesse modelo é particularmente relevante, pois reduz a ocorrência de alarmes incorretos que poderiam levar a paradas desnecessárias de equipamentos e custos adicionais de intervenção.

Analisando o modelo k-NN, percebe-se que, embora o modelo tenha apresentado desempenho similar em termos de acurácia, ele gerou um número maior de falsas falhas. Quando considerados cenários industriais com alta sensibilidade a alarmes falsos, esses resultados podem comprometer a viabilidade prática do modelo. Tem-se assim que os modelos como SVM e XGBoost demonstram desempenho intermediário, sendo adequados quando há restrições computacionais ou necessidade de interpretações mais simples dos resultados. Por outro lado, a Regressão Logística apresentou desempenho inferior, tanto em acurácia quanto em confiabilidade operacional, confirmando sua limitação para problemas com múltiplas classes e relações não lineares.

Assim, é possível concluir que o uso de métricas complementares como o número de falsas falhas, se torna fundamental para a seleção de modelos que sejam aplicáveis em ambientes industriais, pois a acurácia isolada não reflete integralmente a confiabilidade operacional. Além das métricas de desempenho, a interpretação da importância das variáveis permite compreender quais fatores exercem maior influência sobre a previsão de falhas (Figura 6).

**Figura 6.** Importância relativa das variáveis no modelo de Floresta Aleatória, calculada com base no critério de redução de impureza (Gini Importance)




Fonte: Autores.

É possível perceber que os resultados apontam o torque e o desgaste da ferramenta como as variáveis mais importantes, o que mostra como esses fatores são essenciais para prever falhas. O torque está ligado diretamente às condições mecânicas do motor, enquanto o desgaste reflete o tempo de uso e a degradação gradual dos componentes. Também chama atenção a relevância da temperatura do processo, já que o aquecimento excessivo está diretamente relacionado a falhas em motores industriais. A Rotação e temperatura do ar tiveram menor impacto preditivo, embora ainda contribuam para a performance geral do modelo, uma vez que essas variáveis podem interagir indiretamente com as principais.

Essa predominância do torque e do desgaste da ferramenta está alinhada com estudos anteriores, como o de Bezerra et al. (2024), que destacam a importância dessas variáveis para detectar falhas precocemente em sistemas industriais. Isso sugere que estratégias de monitoramento contínuo focadas nesses parâmetros podem aumentar bastante a eficiência da manutenção preditiva. Também, percebe-se que a análise de importância das variáveis também corrobora a decisão de manter todas as variáveis preditoras no modelo. Até mesmo os atributos que apresentaram menor peso relativo podem contribuir de forma indireta para a capacidade preditiva, principalmente em algoritmos de ensemble, que são beneficiados da interação entre múltiplos parâmetros.

## DISCUSSÃO

Percebeu-se que a combinação entre a etapa rigorosa de preparação dos dados e a etapa de aplicação de algoritmos robustos de aprendizado de máquina, como a Floresta Aleatória, pode fornecer previsões altamente confiáveis para a manutenção preditiva de motores industriais. Ainda no modelo de Floresta Aleatória, os resultados mostram a acurácia de 99% e apenas 47 falsas falhas, demonstrando capacidade de generalização e minimizando alarmes incorretos que poderiam gerar intervenções desnecessárias. Tais resultados estão confirmando a robustez dos métodos baseados em ensemble em ambientes industriais complexos, que, conforme destacado por Wu et al. (2024), ressaltam a eficiência dessas técnicas em cenários com múltiplas variáveis correlacionadas e padrões não lineares.

A análise de importância das variáveis (Figura 6) é outro aspecto importante, pois o torque e desgaste da ferramenta foram os principais preditores de falhas, seguidos por temperatura do processo. Achados alinhados com Bezerra et al. (2024), que mostram a relevância das variáveis na previsão de falhas em motores industriais, e com Fioreto et al. (2024) ao destacar a importância de considerar a mitigação de colinearidade e o pré-processamento adequado para maximizar a performance dos modelos preditivos.

Como complemento realizou-se a análise de correlação inicial com Pearson e respectivos p-values. Notou-se que as correlações entre a variável alvo (falha) e os atributos individuais foram fracas ao variar de r = -0,044 a r = 0,191, embora estatisticamente significativas ($p < 0,05$) em todos os casos, com exceção da relação entre torque e desgaste ao apresentar r = -0,003 e p = 0,757. Os valores mostram que, quando analisadas isoladamente, as variáveis não explicam muito bem a ocorrência das falhas. Isso reforça a necessidade de usar modelos que consigam captar relações mais complexas, como as não lineares e as interações entre várias variáveis ao mesmo tempo. Por isso, algoritmos do tipo ensemble, como a Floresta Aleatória, acabam se saindo melhor do que modelos lineares, como a Regressão Logística.

Outro ponto fundamental foi o cuidado com o tratamento dos dados para enfrentar o desbalanceamento entre as classes, um problema muito comum em aplicações reais de aprendizado de máquina. A técnica SMOTE, criada por Chawla et al. (2002), se mostrou eficiente para equilibrar essas classes e aumentar a sensibilidade dos modelos, o que está de acordo com o que Altalhan et al. (2025) destacam ao considerar o SMOTE uma das melhores opções para esse desafio. Além disso, a análise de métricas complementares, como o número de falsas falhas (Tabela 4 e Figura 5), deixa claro que só considerar a acurácia não é suficiente quando vamos escolher modelos para uso industrial. Modelos que têm alta acurácia, mas geram muitos falsos positivos, podem acarretar custos desnecessários e fazer a equipe perder confiança nas previsões. Nesse aspecto, a Floresta Aleatória se destaca por equilibrar bem o desempenho estatístico com a aplicabilidade prática.

Assim, conclui-se que a integração desses resultados com as práticas de TPM reforça o papel estratégico da análise preditiva na gestão de ativos. A antecipação de falhas críticas baseada em dados reais não apenas aumenta a disponibilidade operacional, como também reduz custos de manutenção corretiva e otimiza recursos, consolidando-se como um elemento essencial no contexto da Indústria 4.0.

## CONSIDERAÇÕES FINAIS

Apresentou-se aqui o potencial da integração de técnicas de ML com práticas de TPM na previsão eficaz de falhas em motores industriais. O bom desempenho dos modelos testados, especialmente o da Floresta Aleatória, mostra que ele é uma opção prática e eficiente para ser usado em situações reais de manutenção na indústria.

Foi possível confirmar que algumas variáveis operacionais importantes, como o torque, a temperatura do processo e o desgaste da ferramenta, precisam ser monitoradas o tempo todo para garantir uma operação mais confiável.

Como trabalhos futuros, recomenda-se: (a) a validação prática dos modelos desenvolvidos em ambientes industriais reais; (b) a criação de interfaces integradas a sistemas informatizados para otimizar decisões baseadas nas previsões geradas; (c) a exploração de técnicas avançadas de inteligência artificial, como Redes Neurais Profundas, para buscar ainda melhores resultados preditivos e (d) a automação e integração dos resultados preditivos em Dashboards para melhor visualização e suporte decisório, como em FERREIRA et al., 2024.

Algumas limitações podem se comentadas, pois o estudo utilizou uma base de dados sintética, que pode não refletir integralmente as condições operacionais reais. Outro ponto versa sobre o método de balanceamento adotado (SMOTE), poder induzir o sobre ajuste, principalmente em situações com baixa variabilidade nas classes minoritárias. Tais aspectos precisam ser considerados ao generalizar os resultados para diferentes cenários industriais.

## REFERÊNCIAS


Ahuja, I. P. S. & Khamba, J. S. (2008). Total productive maintenance: literature review and directions. *International Journal of Quality & Reliability Management*, 25(7), 709-56. https://doi.org/10.1108/02656710810890890

UCI Machine Learning Repository. (2020). AI4I 2020 predictive maintenance dataset [Dataset]. *University of California*, Irvine. https://doi.org/10.24432/C5HS5C

Altalhan, M., Algarni, A., & Alouane, M. T. H. (2025). Imbalanced data problem in machine learning: a review. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3531662

Bezerra, F. E., Freitas, C. R. de, Vale, T. F., Alves, M. J., Sales, W. F., & Araújo, A. M. (2024). Impacts of feature selection on predicting machine failures by machine learning algorithms. *Applied Sciences*, 14(8), 3337. https://doi.org/10.3390/app14083337

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-57. https://doi.org/10.1613/jair.953

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37-54. https://doi.org/10.1609/aimag.v17i3.1230

Ferreira, I. T., Freitas, C. R. de, & Luche, J. R. D. (2024). Integração e automação de data lakes com Python para análises em dashboards. *Revista de Gestão e Secretariado,* 15(9), e4105. https://doi.org/10.7769/gesec.v15i9.4105

Fioreto, V. D. L., Freitas, C. R. de, & Luche, J. R. D. (2024). Aplicação de modelos de aprendizado de máquina para a predição da temperatura do rotor em motores PMSM. *Revista de Gestão e Secretariado*, 15(8), e3981. https://doi.org/10.7769/gesec.v15i8.3981

Goyal, C. (2021). Importance of cross validation: Are evaluation metrics enough? Analytics Vidhya. Recuperado de https://www.analyticsvidhya.com/blog/2021/05/importance-of-cross-validation-are-evaluation-metrics-enough/

Kumar, U., Galar, D., Parida, A., Stenström, C., & Berges, L. (2013). Maintenance performance metrics: a state-of-the-art review. *Journal of Quality in Maintenance Engineering*, 19(3), 233-77. https://doi.org/10.1108/JQME-05-2013-0029

Kusiak, A. (2017). Smart manufacturing must embrace big data. *Nature*, 544(7648), 23-5. https://doi.org/10.1038/544023a

Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: a systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, 104, 799-834. https://doi.org/10.1016/j.ymssp.2017.11.016

Mouhib, Z., El Biaze, H., Benabbou, L., & Charkaoui, A. (2025). Total productive maintenance and Industry 4.0: a literature-based path toward a proposed standardized framework. *Applied System Innovation*, 8(4), 98. https://doi.org/10.3390/asi8040098

Muchiri, P. & Pintelon, L. (2008). Performance measurement using overall equipment effectiveness (OEE): literature review and practical application discussion. *International Journal of Production Research*, 46(13), 3517-35. https://doi.org/10.1080/00207540601142645

Neupane, D., Gautam, A., Aryal, A., & Koju, R. (2024). Data-driven machinery fault detection: a comprehensive review. *SSRN Working Paper*. https://doi.org/10.1016/j.neucom.2025.129588

Song, X., Wang, W., Zhang, H., & Zhou, D. (2023). A hybrid deep learning prediction method of remaining useful life for rolling bearings using multiscale stacking deep residual shrinkage network. *International Journal of Intelligent Systems*, 2023(1), 6665534. https://doi.org/10.1155/2023/6665534

Tortorella, G., Sawhney, R., Jurburg, D., & Fogliatto, F. S. (2022). The impact of Industry 4.0 on the relationship between TPM and maintenance performance. *Journal of Manufacturing Technology Management*, 33(3), 489520. https://doi.org/10.1108/JMTM-10-2021-0399

Wu, M., Li, X., Sun, J., Wang, Z., & Hu, Y. (2024). An intelligent predictive maintenance system based on random forest for addressing industrial conveyor belt challenges. *Frontiers in Mechanical Engineering*, 10, 1383202. https://doi.org/10.3389/fmech.2024.1383202

Zonta, T., da Costa, C. A., Righi, R. da R., Lima, M. J. de, Trindade, E. S. da, & Li, G. P. (2020). Predictive maintenance in the industry 4.0: a systematic literature review. *Computers & Industrial Engineering*, 150, 106889. https://doi.org/10.1016/j.cie.2020.106889