

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE  
MINAS GERAIS - *CAMPUS* IBIRITÉ  
ENGENHARIA DE CONTROLE E AUTOMAÇÃO

Victor Henrique de Mattos Medeiros

**INTEGRAÇÃO DA INTELIGÊNCIA ARTIFICIAL NA GESTÃO DE  
ATIVOS DE TECNOLOGIA DA AUTOMAÇÃO**

Ibirité - MG  
2025

VICTOR HENRIQUE DE MATTOS MEDEIROS

**INTEGRAÇÃO DA INTELIGÊNCIA ARTIFICIAL NA GESTÃO DE  
ATIVOS DE TECNOLOGIA DA AUTOMAÇÃO**

Trabalho de conclusão de curso apresentado ao Curso de Engenharia de Controle e Automação do Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais - *Campus* Ibirité para a obtenção do título de Engenheiro de Controle e Automação.

**Orientador:** Prof. Dr. Diego Henrique de Souza Chaves

Ibirité - MG  
2025

---

M488i  
2025

Medeiros, Victor Henrique de Mattos.

Integração da inteligência artificial na gestão de ativos de tecnologia da automação. [manuscrito] / Victor Henrique de Mattos Medeiros. – Ibirité: Instituto Federal de Minas Gerais, 2025.

55 f.; il.

Trabalho de Conclusão de Curso (Bacharelado em Engenharia de controle e automação) – Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais – Campus Ibirité, 2025.

Orientador: Prof. Dr. Diego Henrique de Souza Chaves.

1. Inteligência artificial. 2. Cluster (sistema de computador). 3. Gestão de ativos. I. Chaves, Diego Henrique de Souza. II. Instituto Federal de Minas Gerais – *Campus* Ibirité. III. Título

CDU 004.8

---


Catálogo: Luciana Batista Neves - CRB-6/2000

Victor Henrique de Mattos Medeiros

## **INTEGRAÇÃO DA INTELIGÊNCIA ARTIFICIAL NA GESTÃO DE ATIVOS DE TECNOLOGIA DA AUTOMAÇÃO**


Trabalho de conclusão de curso apresentado ao Curso de Engenharia de Controle e Automação do Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais - *Campus* Ibirité para a obtenção do título de Engenheiro de Controle e Automação.

Aprovado em: 27/02/2025 pela banca examinadora:

Documento assinado digitalmente  
 **DIEGO HENRIQUE DE SOUZA CHAVES**  
Data: 03/04/2025 15:38:35-0300  
Verifique em <https://validar.iti.gov.br>


---

Prof. Dr. Diego Henrique de Souza Chaves - IFMG (Orientador)

Documento assinado digitalmente  
 **CARLOS DIAS DA SILVA JUNIOR**  
Data: 03/04/2025 12:48:37-0300  
Verifique em <https://validar.iti.gov.br>

---

Prof. Me. Carlos Dias da Silva Junior - IFMG

Documento assinado digitalmente  
 **THIAGO HENRIQUE BARBOSA DE CARVALHO TA**  
Data: 02/04/2025 23:13:50-0300  
Verifique em <https://validar.iti.gov.br>

---

Prof. Dr. Thiago Henrique Barbosa de Carvalho Tavares - IFMG

Dedico esta monografia aos meus amados pais, maiores incentivadores e fontes inesgotáveis de apoio, amor e compreensão.

## **AGRADECIMENTOS**

Agradeço a Deus, meu Pai, maravilhoso conselheiro que é o motivo pelo qual me dedico e busco fazer a diferença no mundo.

Agradeço a toda à minha família por acreditarem em mim e pelo incentivo constante durante toda a minha vida, sendo essenciais para que eu pudesse me desenvolver como homem. Gostaria de expressar minha gratidão à minha namorada, Kemilly, que me ajudou em todos os momentos.

Agradeço ao meu orientador, Diego Henrique de Souza Chaves, que foi extremamente competente em me auxiliar em todos os desafios e dúvidas que surgiram. Um agradecimento especial ao professor Pedro Augusto Ramos, ao qual juntos trabalhamos em dois projetos de pesquisa, que fez com que eu desenvolvesse vários aspectos da minha vida acadêmica e ajudou em meu crescimento pessoal.

“Lembre sempre daquilo que aprendeu. A sua educação é a sua vida; guarde-a bem.”

Provérbios 4.13

## RESUMO

A gestão de ativos de tecnologia da automação é uma prática necessária na eficiência operacional das indústrias. No entanto, um dos grandes desafios enfrentados pelas empresas é a falta de um sistema estruturado para gestão de ativos com base em fatores críticos e operacionais, como risco operacional, ciclo de vida, sobressalentes e impacto financeiro. Decisões tomadas sem um critério bem definido, pode levar a falhas inesperadas, desperdício de recursos e intervenções de manutenção ineficientes. Este trabalho propõe a integração da Inteligência Artificial (IA) para auxiliar na tomada de decisão na gestão de ativos, utilizando técnicas de aprendizado de máquina e clusterização. A metodologia adotada incluiu a coleta e tratamento de dados de ativos de automação de uma planta industrial real, a definição de um critério de priorização baseado em múltiplas variáveis e a aplicação do algoritmo K-Means para segmentação dos ativos em clusters. A qualidade da clusterização foi avaliada por métricas como Silhouette Score, garantindo uma divisão dos ativos conforme sua criticidade e relevância operacional. Os resultados demonstraram que a abordagem proposta permite categorizar ativos de forma a identificar aqueles com maior impacto operacional e necessidade de atenção prioritária. Além disso, foi desenvolvida uma interface interativa para visualização dos clusters, facilitando a interpretação dos dados e auxiliando na tomada de decisão estratégica. Os resultados demonstram que a utilização de IA na gestão de ativos pode contribuir para a redução de custos, a mitigação de riscos e o aumento da confiabilidade operacional, representando um avanço significativo para a manutenção preditiva e gestão estratégica de ativos na indústria.

**Palavras-chave:** Inteligência Artificial. Gestão de Ativos. Clusterização.

## ABSTRACT

The management of automation technology assets is essential for ensuring the operational efficiency of industries. However, one of the major challenges faced by companies is the lack of a structured asset management system based on critical and operational factors such as operational risk, lifecycle, spare parts availability, and financial impact. Decisions made without a well-defined criterion can lead to unexpected failures, resource waste, and inefficient maintenance interventions. This study proposes the integration of Artificial Intelligence (AI) to support decision-making in asset management by employing machine learning techniques and clustering. The adopted methodology included data collection and processing of automation assets from a real industrial plant, the definition of a prioritization criterion based on multiple variables, and the application of the K-Means algorithm to segment assets into clusters. The clustering quality was assessed using metrics such as the Silhouette Score, ensuring that assets were grouped according to their criticality and operational relevance. The results demonstrated that the proposed approach effectively categorizes assets, identifying those with higher operational impact and requiring prioritized attention. Additionally, an interactive interface was developed to visualize the clusters, facilitating data interpretation and supporting strategic decision-making. The findings indicate that the use of AI in asset management can contribute to cost reduction, risk mitigation, and increased operational reliability, representing a significant advancement in predictive maintenance and strategic asset management within the industry.

**Keywords:** Artificial Intelligence. Asset Management. Clustering.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Relação entre Inteligência Artificial, Machine Learning e Clusterização. . .	19
Figura 2 – Funcionamento do K-Means. . . . .	20
Figura 3 – Ciclo de vida dos ativos. . . . .	26
Figura 4 – Trecho do código em Jupyter Notebook. . . . .	27
Figura 5 – Exemplificação do Método do Cotovelo. . . . .	30
Figura 6 – Exemplificação do Método K-Means. . . . .	31
Figura 7 – Fluxograma do funcionamento do algoritmo. . . . .	33
Figura 8 – Relação entre Prioridade e Risco Total. . . . .	37
Figura 9 – Relação entre Prioridade e Criticidade. . . . .	38
Figura 10 – Gráfico do Método do Cotovelo. . . . .	39
Figura 11 – Distribuição de Ativos por Cluster. . . . .	40
Figura 12 – Prioridade média em relação a cada cluster. . . . .	42
Figura 13 – Distribuição do Ciclo de Vida dos Ativos em cada Cluster. . . . .	43
Figura 14 – Criticidade Média por Cluster. . . . .	44
Figura 15 – Distribuição da Relacao_Instalada_Sobressalente em função da Prioridade Calculada por Cluster. . . . .	45
Figura 16 – Histograma da distribuição do Risco Total por Cluster. . . . .	46
Figura 17 – Valor médio dos Ativos por Cluster. . . . .	47
Figura 18 – Histograma da distribuição do Valor dos Ativos por Cluster. . . . .	48
Figura 19 – Arquivo Excel exportado pelo código desenvolvido. . . . .	49
Figura 20 – Exemplo de interação com a interface. . . . .	49

## LISTA DE TABELAS

Tabela 1 – Descrição das Variáveis Relacionadas à Gestão de Ativos. . . . .	25
Tabela 2 – Descrição das Variáveis Relacionadas à Gestão de Riscos. . . . .	26
Tabela 3 – Mapeamento das Variáveis Categóricas para Numéricas. . . . .	28
Tabela 4 – Principais variáveis analisadas. . . . .	34
Tabela 5 – Estatísticas descritivas antes da normalização. . . . .	35
Tabela 6 – Estatísticas descritivas após a normalização. . . . .	35
Tabela 7 – Estatísticas descritivas dos ativos do Cluster 0. . . . .	41
Tabela 8 – Estatísticas descritivas dos ativos do Cluster 1. . . . .	41
Tabela 9 – Estatísticas descritivas dos ativos do Cluster 2. . . . .	41

## **LISTA DE ABREVIATURAS E SIGLAS**

IA	Inteligência Artificial
ML	Machine Learning
NLP	Natural Language Processing
TA	Tecnologia de Automação
CLP	Controlador Lógico Programável
SCADA	Supervisory Control and Data Acquisition
IoT	Internet of Things
CPS	Sistemas Ciber Físicos
RFID	Radio Frequency Identification
AR	Realidade Aumentada
TI	Tecnologia da Informação
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
CBM	Condition-Based Maintenance
Pandas	Python Data Analysis Library
WSS	Within-Cluster Sum of Squares
SSE	Somatório dos erros quadráticos dos clusters
CAPEX	Capital Expenditure

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>12</b>
<b>1.1</b>	<b>Objetivos</b>	<b>13</b>
<i>1.1.1</i>	<i>Objetivo geral</i>	<i>13</i>
<i>1.1.2</i>	<i>Objetivos específicos</i>	<i>14</i>
<b>1.2</b>	<b>Justificativa</b>	<b>14</b>
<b>1.3</b>	<b>Organização do Texto</b>	<b>15</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>16</b>
<b>2.1</b>	<b>Desafios da Indústria 4.0</b>	<b>16</b>
<b>2.2</b>	<b>Manutenção Preditiva na Indústria 4.0</b>	<b>16</b>
<b>2.3</b>	<b>Inteligência Artificial</b>	<b>17</b>
<i>2.3.1</i>	<i>Machine Learning</i>	<i>17</i>
<i>2.3.2</i>	<i>Clusterização</i>	<i>18</i>
<i>2.3.3</i>	<i>K-Means</i>	<i>20</i>
<b>2.4</b>	<b>Integração da Inteligência Artificial na Gestão de Ativos</b>	<b>21</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>25</b>
<b>3.1</b>	<b>Base de Dados</b>	<b>25</b>
<b>3.2</b>	<b>Tratamento dos Dados</b>	<b>26</b>
<i>3.2.1</i>	<i>Organização</i>	<i>26</i>
<i>3.2.2</i>	<i>Tratamento</i>	<i>27</i>
<b>3.3</b>	<b>Definição e Cálculo do Critério de Prioridade</b>	<b>28</b>
<b>3.4</b>	<b>Clusterização</b>	<b>30</b>
<b>3.5</b>	<b>Interface para Exploração dos Clusters</b>	<b>31</b>
<b>4</b>	<b>RESULTADOS</b>	<b>34</b>
<b>4.1</b>	<b>Caracterização dos Dados e Pré-Processamento</b>	<b>34</b>
<b>4.2</b>	<b>Cálculo do Critério de Prioridade dos Ativos</b>	<b>36</b>
<b>4.3</b>	<b>Identificação dos Clusters de Ativos de Automação</b>	<b>38</b>
<b>4.4</b>	<b>Caracterização dos Clusters</b>	<b>39</b>
<b>4.5</b>	<b>Análise dos Clusters e Relevância para a Gestão de Ativos</b>	<b>42</b>
<b>4.6</b>	<b>Exportação dos Dados e Aplicação Prática</b>	<b>48</b>

<b>5</b>	<b>CONCLUSÃO E TRABALHOS FUTUROS . . . . .</b>	<b>51</b>
<b>5.1</b>	<b>Trabalhos Futuros . . . . .</b>	<b>52</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>53</b>

# 1 INTRODUÇÃO

A Inteligência Artificial surgiu em meados do Século XX com o intuito de criar máquinas que pudessem replicar as capacidades humanas de raciocínio e aprendizado. O termo “Inteligência Artificial” foi introduzido por John McCarthy em 1956, durante a Dartmouth Summer Research Project on Artificial Intelligence, uma conferência que é amplamente considerada o ponto de partida da IA (McCarthy *et al.*, 2006).

Posteriormente, houveram momentos em que os investimentos e a confiança na IA cresceram exponencialmente, impulsionados por promessas de revoluções tecnológicas, mas esses momentos foram seguidos por um período de desilusão e redução de recursos (Sichman, 2021), período conhecido como os “Invernos da IA”.

O primeiro Inverno da IA ocorreu na década de 1970, seguido por outro na década de 1980. A limitação das tecnologias disponíveis na época, aliada à dificuldade em resolver problemas complexos de forma eficiente, resultou em um cenário de frustração na comunidade científica e empresarial (Hendler, 2008). Como consequência, aconteceram cortes significativos nos investimentos e redução da pesquisa no campo da Inteligência Artificial, adiando por décadas alguns dos avanços que só seriam retomados com o progresso da capacidade computacional e novas abordagens algorítmicas.

Ademais, o campo da Inteligência Artificial superou esse período ao redefinir expectativas, concentrando em resolução de problemas lógicos e manejáveis, além de estar em concordância com os avanços tecnológicos, que facilitaram o processamento dos dados e execução de algoritmos (Hollan; Hutchins; Kirsh, 2000).

Com o tempo, avanços tecnológicos permitiram o desenvolvimento de algoritmos mais complexos, com capacidade de processar grandes volumes de dados. Desta forma, fez com que a IA fosse levada a expandir-se para áreas como Machine Learning (ML), deep learning, visão computacional e processamento de linguagem natural (NLP - Natural Language Processing). Atualmente a IA está difundida e integrada em diversas aplicações práticas, auxiliando e sendo fundamental em setores como indústria, saúde, finanças e transporte (Russell; Norvig, 2016).

Dentro da Inteligência Artificial, há subcampos, como Machine Learning (ML), que se concentra no desenvolvimento de algoritmos e técnicas que possibilitam que computadores aprendam e façam previsões ou decisões baseadas em dados. Desta forma, os algoritmos podem ser aplicados na identificação de padrões em dados complexos e heterogêneos, fornecendo uma base sólida para diversas aplicações industriais (Bishop; Nasrabadi, 2006).

A clusterização é um método de aprendizado não supervisionado, sendo assim não há a presença de algo que indique ao sistema o significado de cada padrão presente nos dados (Jain; Dubes, 1988). Em vez disso, o algoritmo trabalha identificando similaridades entre os elementos, agrupando-os com base em suas características intrínsecas. A previsão ou agrupamento é realizado a partir da representação interna construída pelo sistema, utilizando exclusivamente o

conjunto de dados fornecido como entrada.

Na manutenção preditiva, a IA é capaz de analisar grandes volumes de dados, que podem ser provenientes de sensores instalados em equipamentos industriais. Por meio do aprendizado de máquina, a IA pode criar modelos preditivos baseados em dados históricos e em tempo real. Esses permitem a previsão antes que a falha ocorra (Carvalho *et al.*, 2019). Contudo, a IA também pode ser utilizada na gestão de ativos de Tecnologia de Automação (TA).

Os ativos de automação são os recursos físicos, tecnológicos e digitais utilizados para automatizar processos industriais, comerciais ou de serviços. Esses ativos englobam desde hardware e software até infraestruturas de comunicação que permitem o controle e monitoramento de processos sem a necessidade de intervenção humana constante, como CLPs (Controladores Lógicos Programáveis), sistemas SCADA (Supervisory Control and Data Acquisition) ou switches industriais.

A gestão de ativos de TA é um conjunto de práticas que envolvem a supervisão estratégica e sistemática de todos os componentes tecnológicos utilizados na automação de processos industriais (EMERICK, 2019). A gestão de ativos permite a avaliação contínua do estado dos ativos, permitindo que decisões sejam tomadas de forma mais eficiente e planejada (Minnaar; Basson; Vlok, 2013).

Com integração em IA, a gestão de ativos representa uma evolução do gerenciamento dos processos industriais, pois oferece novos níveis de eficiência, previsibilidade, controle e otimização (Group, 2024). Essa integração demonstra como a gestão de ativos em ambientes industriais pode ser transformada, utilizando algoritmos de aprendizado de máquinas e análise de dados.

Contudo, a gestão de ativos está em estágio inicial no debate científico e são necessárias soluções que apoiem sua adoção dentro da indústria (Paul; Ogugua; Eyo-Udo, 2024). Neste trabalho, é explorado como a IA é capaz de transformar a gestão de ativos, proporcionando previsibilidade, confiabilidade e otimização dos processos industriais na área da automação.

## 1.1 Objetivos

### 1.1.1 *Objetivo geral*

O objetivo geral do trabalho é desenvolver um modelo baseado em inteligência artificial para a gestão de ativos de tecnologia da automação. O trabalho integra a análise de dados e técnicas de clusterização para auxiliar na tomada de decisão na gestão de ativos, com o objetivo de melhorar a eficiência da gestão e reduzir os riscos operacionais em uma planta industrial real.

### 1.1.2 *Objetivos específicos*

I- Definir as informações do banco de dados a serem utilizadas para a realização do processo de limpeza e tratamento dos dados;

II- Escolher a linguagem de programação a ser utilizada para desenvolver o algoritmo e processar os dados;

III- Importar a base de dados com as informações mais relevantes para auxiliar o desenvolvimento, cálculo de parâmetros, processamento e refinamento dos dados;

IV- Escolher os métodos de tratamento dos dados conforme a necessidade e os tipos de dados apresentados na base de dados referente a gestão dos ativos e gestão de riscos dos equipamentos;

V- Definição da forma de cálculo do parâmetro que auxiliará a tomada de decisão referente a gestão de ativos;

VI- Aplicar a técnica de aprendizado não supervisionado de clusterização, utilizando os parâmetros fornecidos e calculados;

VII- Utilizar os dados dos agrupamentos de cada cluster com o intuito de avaliar o desempenho da IA;

VIII- Exportação dos dados gerados que auxilia na tomada de decisão da gestão de ativos a fim de permitir uma visualização aprofundada;

## 1.2 **Justificativa**

Mesmo diante de toda a evolução tecnológica e acessibilidade à informação, enfrentamos diversos problemas e limitações. Por isso, é necessário gerir os ativos de maneira que impacte diretamente na eficiência operacional e nos custos. A abordagem tradicional de gestão, baseada em planilhas manuais e métodos estatísticos apresenta limitações, como suscetibilidade a erros humanos, dificuldade em acompanhar atualizações em tempo real e incapacidade de processar análises complexas que envolvam grandes volumes de dados.

Dessa forma, o uso de técnicas de inteligência artificial (IA) surge como uma alternativa para transformar o modelo atual em um sistema dinâmico e automatizado. A clusterização como técnica de aprendizado não supervisionado permite categorizar os ativos com base em critérios como ciclo de vida, criticidade, probabilidade de falha e impacto operacional, o que possibilita a priorização de ações de manutenção e substituição a fim de tratar a obsolescência.

Por isso, esse trabalho integra a inteligência artificial à gestão de ativos de maneira que proporcione maior eficácia na tomada de decisões, assim como na redução de riscos operacionais e otimização de custos. Essa abordagem promove a modernização dos processos de gestão, alinhando com as demandas oriundas da Indústria 4.0, onde a automação e a análise de dados

são pilares. Desse modo, ao utilizar dados reais de uma planta industrial, o projeto apresenta não somente valor acadêmico, mas também impacto direto na melhoria dos processos de gestão, controle de processos e competitividade empresarial.

### 1.3 Organização do Texto

Este trabalho está organizado da seguinte forma:

#### (1) Introdução

Na introdução, é apresentada uma breve contextualização sobre temas fundamentais do trabalho, como também expõe algumas das necessidades que motivaram o desenvolvimento do trabalho. Além disso, são encontrados o objetivo geral, os objetivos específicos, a justificativa e a disposição da estrutura do trabalho.

#### (2) Revisão Bibliográfica

Na revisão bibliográfica são apresentadas diversas literaturas e trabalhos que possuem temas similares às áreas que são pilares neste trabalho. Desta forma, são fonte de estudo para o desenvolvimento do mesmo.

#### (3) Metodologia

No terceiro capítulo, são descritos os métodos utilizados para a obtenção, tratamento e análise dos dados e do cálculo do critério estabelecido. Também são apresentados os métodos de aprendizado não supervisionado utilizados, assim como os métodos de análise dos resultados obtidos.

#### (4) Resultados

No quarto capítulo, são mostrados os gráficos, tabelas e imagens referentes aos resultados adquiridos com a técnica de aprendizado não supervisionado. Sendo assim, os resultados foram analisados explorando as contribuições do trabalho.

#### (5) Conclusão

No último capítulo, foram apresentadas as conclusões finais do trabalho, um resumo e uma visão geral sobre o modelo obtido, bem como a importância do tema e os possíveis trabalhos futuros.

#### (6) Referências

Ao final, foram incluídas referências de artigos, livros, documentos, que contribuíram de para o desenvolvimento deste trabalho.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 Desafios da Indústria 4.0

A Indústria 4.0 representa uma transformação nos processos produtivos impulsionada por tecnologias como Internet das Coisas (IoT), Sistemas Ciber-Físicos (CPS), Inteligência Artificial (IA) e Computação em Nuvem. Essas inovações possibilitam a comunicação autônoma entre dispositivos, a monitorização em tempo real e a criação de fábricas inteligentes (Santos *et al.*, 2018).

No entanto, a adoção da Indústria 4.0 apresenta desafios, como a padronização das interfaces de comunicação, a segurança digital e a reestruturação dos processos de trabalho para acompanhar essas mudanças. Tecnologias emergentes, como RFID, manufatura aditiva (impressão 3D) e realidade aumentada (AR), facilitam a gestão de estoques, a personalização da produção e a interação homem-máquina (Santos *et al.*, 2018).

Além das mudanças tecnológicas, a Indústria 4.0 tem impactos diretos na organização do trabalho. Segundo Tessarini (2018), a automação reduz postos de trabalho menos qualificados, mas cria novas oportunidades que exigem habilidades avançadas em TI, análise de dados e interação com sistemas inteligentes. Isso exige a requalificação dos trabalhadores e a adaptação das organizações para esse novo modelo produtivo, que se torna mais dinâmico, colaborativo e digitalizado.

A integração digital dentro das empresas facilita a troca de informações em tempo real entre diferentes setores e organizações, o que aumenta a eficiência, competitividade e flexibilidade. No entanto, sua implementação requer cuidados estratégicos, incluindo investimentos em segurança digital e capacitação da mão de obra para lidar com as novas demandas do mercado (Tropia; Silva; Dias, 2017).

### 2.2 Manutenção Preditiva na Indústria 4.0

A manutenção preditiva tem se consolidado como uma estratégia que promove a alta disponibilidade e eficiência dos equipamentos industriais na Indústria 4.0. Diferente das abordagens tradicionais, essa metodologia permite a antecipação de falhas por meio da análise contínua de dados coletados por sensores inteligentes. À medida que os processos produtivos exigem máxima operacionalidade, a manutenção passou a desempenhar um papel estratégico, reduzindo custos e minimizando paradas não programadas (Baldissarelli; Fabro, 2019).

Um exemplo prático dessa abordagem foi a implementação de um sistema de manutenção preditiva em um exaustor industrial, onde sensores de vibração monitoram em tempo real o estado dos rolamentos e do motor. Essa solução permitiu a detecção remota de anomalias, possibilitando intervenções preventivas antes que falhas críticas ocorressem. Além da redução dos custos com reparos emergenciais, houve um aumento significativo na disponibilidade do equi-

pamento e na eficiência da equipe de manutenção, que passou a focar em atividades de maior valor agregado (Baldissarelli; Fabro, 2019).

A aplicação da manutenção preditiva também se estende a sistemas críticos de computação, onde a interrupção de serviços pode gerar impactos significativos. A crescente demanda por infraestruturas de computação de alto desempenho (HPC) e serviços em nuvem tornou essa abordagem indispensável (Lima; Aranha; Sperandio, 2021). Ambientes de missão crítica, compostos por milhares de nós computacionais interconectados, operam ininterruptamente e exigem infraestrutura redundante para garantir sua confiabilidade. Qualquer falha pode comprometer a disponibilidade dos sistemas e elevar consideravelmente os custos operacionais.

## 2.3 Inteligência Artificial

A Inteligência Artificial (IA) é um campo da ciência da computação que visa desenvolver sistemas capazes de realizar tarefas que normalmente exigem inteligência humana, como aprendizado, raciocínio, percepção e tomada de decisão (Russell; Norvig, 2016). O termo "Inteligência Artificial" foi introduzido por John McCarthy em 1956, durante a Dartmouth Summer Research Project on Artificial Intelligence, evento considerado o marco inicial da IA como área de pesquisa formal (McCarthy *et al.*, 2006).

Desde então, a IA evoluiu significativamente, impulsionada por avanços em algoritmos, maior capacidade de processamento e a disponibilidade de grandes volumes de dados. Atualmente, a IA abrange diversas subáreas, incluindo aprendizado de máquina (Machine Learning - ML), redes neurais artificiais, processamento de linguagem natural (Natural Language Processing - NLP) e visão computacional (Sichman, 2021). O aprendizado de máquina, em particular, tem se destacado como uma abordagem que viabiliza o desenvolvimento de sistemas inteligentes, para que algoritmos ajustem seus modelos a partir de padrões extraídos de grandes conjuntos de dados (Bishop; Nasrabadi, 2006).

### 2.3.1 Machine Learning

Machine Learning (ML), ou Aprendizado de Máquina, é um ramo da Inteligência Artificial (IA) que permite que sistemas aprendam a partir de dados e melhorem seu desempenho sem a necessidade de programação explícita (Russell; Norvig, 2016). Esse conceito surgiu como um desdobramento dos estudos de IA propostos por McCarthy (McCarthy *et al.*, 2006), que buscavam criar máquinas capazes de raciocinar e tomar decisões autônomas.

O ML impulsiona a automação inteligente e a otimização de processos (Santos *et al.*, 2018). As aplicações incluem desde manutenção preditiva em fábricas até análises avançadas de dados industriais para melhorar a eficiência operacional (Raptis; Passarella; Conti, 2019). No entanto, o uso crescente de IA e ML também traz desafios, incluindo a necessidade de garantir transparência nos modelos e evitar viés algorítmico (Hendler, 2008).

O aprendizado de máquina pode ser classificado em diferentes categorias, sendo que as principais são: aprendizado supervisionado e aprendizado não supervisionado.

No aprendizado supervisionado, o modelo é treinado com um conjunto de dados rotulado, onde cada entrada possui uma saída correspondente conhecida. O objetivo do algoritmo é identificar padrões nos dados e generalizar esse conhecimento para prever novas instâncias (Bishop; Nasrabadi, 2006).

Esse método é amplamente utilizado em tarefas como reconhecimento de voz, diagnóstico médico e detecção de fraudes financeiras. Por exemplo, um modelo de aprendizado supervisionado pode ser treinado com dados históricos de transações bancárias, aprendendo a diferenciar padrões normais de comportamentos fraudulentos para alertar sobre possíveis fraudes (Raptis; Passarella; Conti, 2019).

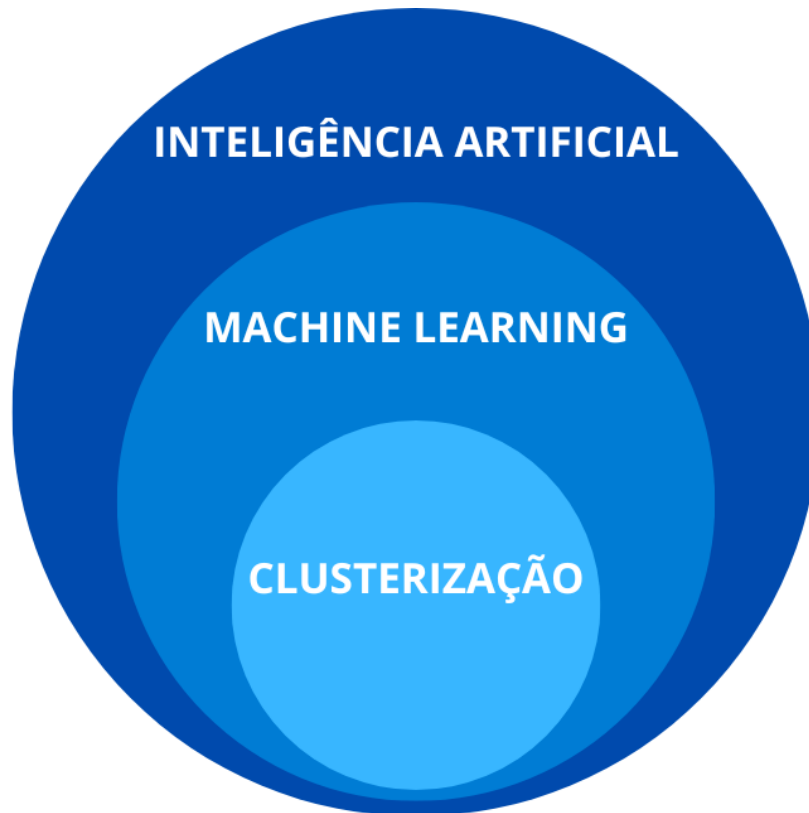
No aprendizado não supervisionado, os dados de entrada não possuem rótulos, sendo assim, não possuem saídas conhecidas e o modelo deve encontrar padrões ou agrupamentos ocultos sem supervisão explícita (Bishop; Nasrabadi, 2006). Esse método é útil para descoberta de conhecimento, onde padrões podem emergir sem um conhecimento prévio sobre as categorias existentes nos dados.

As aplicações desta técnica incluem segmentação de clientes, compressão de dados e análise de tendências de mercado. Um exemplo é a análise de clusters, onde empresas utilizam algoritmos para agrupar consumidores com comportamentos similares, permitindo estratégias de marketing personalizadas (Sichman, 2021).

### **2.3.2 Clusterização**

A clusterização é uma técnica de grande relevância na mineração de dados e em análises exploratórias, permitindo descobrir grupos e padrões que estão ocultos nos dados (Jain; Dubes, 1988). Essa técnica pode ser aplicada em diversos contextos que exigem identificação de padrões, a fim de otimizar o planejamento. Métodos como K-means, hierárquico e DBSCAN (Density-Based Spatial Clustering of Applications with Noise) são exemplos de algoritmos frequentemente utilizados na clusterização.

Figura 1 – Relação entre Inteligência Artificial, Machine Learning e Clusterização.



Fonte: Elaborado pelo autor, 2025.

A clusterização é uma técnica focada em agrupar elementos de um conjunto de dados similares no mesmo cluster, enquanto objetos menos similares são alocados em clusters diferentes (Ochi; Dias; Soares, 2004). Os autores destacam que a clusterização pode ser classificada em dois tipos principais: a clusterização com número de grupos definidos (K-clusterização) e a clusterização automática, na qual o número de clusters é determinado durante o processo.

Os autores exploram diferentes métodos de clusterização, que incluem algoritmos hierárquicos, particionados e baseados em densidade. Entre os algoritmos discutidos, destacam-se o K-Means, que é amplamente utilizado devido à sua eficiência em grandes volumes de dados, e o DBSCAN, que se destaca por lidar melhor com clusters de formatos variados e a presença de ruído nos dados.

O estudo também enfatiza a importância da clusterização na análise de grandes bases de dados, demonstrando sua aplicação em diversas áreas, como reconhecimento de padrões, bioinformática e sistemas de recomendação. Além disso, os autores discutem métricas utilizadas para avaliar a qualidade dos agrupamentos, como coeficiente de silhueta (silhouette score) e índices de validade (Davies-Bouldin Index). O trabalho contribui para a compreensão da clusterização como ferramenta analítica, destacando sua aplicabilidade na tomada de decisões estratégicas em diversos setores.

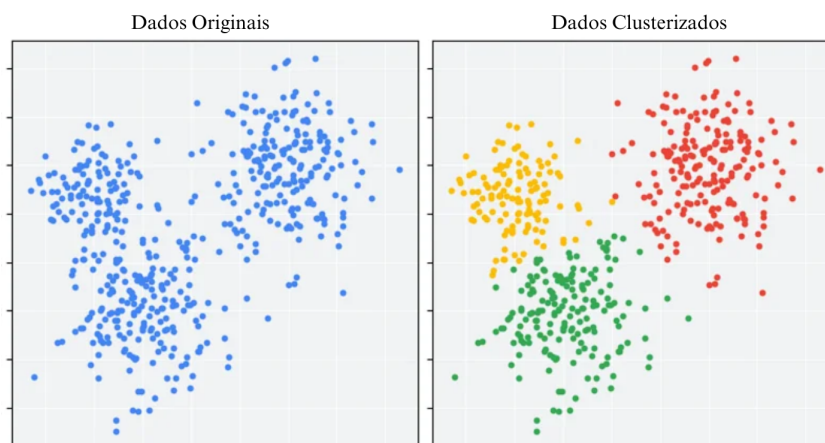
### 2.3.3 K-Means

O *K-Means* é uma técnica amplamente utilizada para clusterização de dados, sendo classificado como um método de aprendizado não supervisionado. Seu objetivo é particionar um conjunto de dados em  $k$  grupos distintos, garantindo que os elementos dentro de um mesmo cluster sejam mais semelhantes entre si do que em relação a elementos de outros clusters. O *K-Means* destaca-se por sua eficiência computacional e aplicabilidade em diversos domínios, como análise de mercado, detecção de padrões e manutenção preditiva (Zubair *et al.*, 2024).

O algoritmo *K-Means* opera iterativamente, seguindo os seguintes passos:

1. Definição do número de clusters  $k$ .
2. Seleção de  $k$  pontos iniciais como centróides, geralmente escolhidos de forma aleatória.
3. Atribuição de cada ponto do conjunto de dados ao cluster cujo centróide esteja mais próximo, utilizando uma métrica de distância, como a distância euclidiana.
4. Atualização dos centróides recalculando a média dos pontos atribuídos a cada cluster.
5. Repetição dos passos de realocação dos pontos e recalculação dos centróides até que a convergência seja atingida, ou seja, até que os centróides não mudem significativamente entre as iterações.

Figura 2 – Funcionamento do K-Means.



Fonte: Adaptado de Poli Júnior, 2024.

A escolha do número ideal de clusters ( $k$ ) é um fator crítico para a efetividade do *K-Means*. Diferentes abordagens podem ser utilizadas para essa definição, sendo o Método do Cotovelo (*Elbow Method*) e o *Silhouette Score* as mais comuns. O Método do Cotovelo analisa a soma dos erros quadráticos dentro dos clusters (*SSE - Sum of Squared Errors*) e determina o valor de  $k$  com base no ponto de inflexão da curva, onde a adição de mais clusters deixa de

reduzir significativamente o erro (Bholowalia; Kumar, 2014). Já o Índice de Silhouette mede a separabilidade dos clusters, avaliando o quão próximo um ponto está de seu próprio cluster em comparação com outros. Valores próximos de 1 indicam clusters bem definidos, enquanto valores negativos sugerem uma alocação inadequada dos pontos (Vardakas; Papakostas; Likas, 2024).

O *K-Means* possui diversas aplicações em setores industriais e comerciais. Um dos usos mais comuns é na segmentação de clientes, onde empresas agrupam consumidores com comportamentos de compra semelhantes para personalizar estratégias de marketing (Lopes; Cavique, 2022). Na indústria, o algoritmo pode ser empregado para manutenção preditiva, permitindo a categorização de equipamentos de acordo com seu risco operacional e auxiliando na priorização de intervenções de manutenção (Costa *et al.*, 2024). Além disso, o *K-Means* é amplamente utilizado na detecção de anomalias, sendo uma ferramenta fundamental para identificar padrões incomuns em segurança cibernética e prevenção de fraudes financeiras.

A utilização do *K-Means* na gestão de ativos industriais representa uma abordagem promissora, pois possibilita a categorização de equipamentos com base em múltiplos fatores críticos, contribuindo para a tomada de decisões estratégicas e para a otimização de recursos. Sua capacidade de agrupar ativos com características semelhantes permite um planejamento mais eficiente da manutenção e do estoque de peças sobressalentes, garantindo maior confiabilidade operacional e redução de custos.

## 2.4 Integração da Inteligência Artificial na Gestão de Ativos

A integração de técnicas de Inteligência Artificial (IA), como o aprendizado profundo (Deep Learning), auxiliam a manipulação de grandes volumes de dados (Big Data), permitindo a identificação de padrões complexos e a previsão de falhas com alta precisão, que resulta em um suporte mais eficiente para a tomada de decisões (Carvalho *et al.*, 2019). Dessa forma, as manutenções se tornam planejadas e são realizadas de forma a reduzir o tempo de inatividade e os custos de operação, enquanto aumenta a confiabilidade e o desempenho do ambiente.

Com a crescente adoção de tecnologias fundamentadas em Inteligência Artificial, a análise de grandes volumes de dados industriais tornou-se foco para a otimização de processos e gestão eficiente. Uma técnica amplamente utilizada nesse contexto é a clusterização, que consiste em identificar padrões e agrupar elementos semelhantes em conjuntos (clusters), a partir de características comuns. Esse método permite explorar grandes bases de dados, sem a necessidade de que algo seja indicado no sistema (Oyewole; Thopil, 2023).

A clusterização tem se mostrado uma ferramenta valiosa na gestão de ativos industriais, permitindo a identificação de padrões operacionais e a segmentação de equipamentos com base em seu comportamento. No estudo de Pinto (2021), essa técnica foi aplicada na manutenção preditiva de válvulas industriais, utilizando o algoritmo K-Means para analisar parâmetros como

frequência de acionamento, tempo de operação e histórico de falhas.

A abordagem utilizada no estudo possibilitou a formação de grupos de válvulas com características similares, o que permitiu um monitoramento mais eficiente e tomadas de decisão mais assertivas na gestão da manutenção. Além de contribuir para a previsão de falhas, a clusterização permitiu priorizar intervenções e distribuir melhor os recursos, otimizando estoques de peças de reposição e planejamento de manutenção preventiva.

Outro ponto destacado no estudo é a integração da clusterização com ferramentas de análise visual, permitindo que os clusters identificados fossem monitorados em tempo real por engenheiros e operadores. Isso facilitou a rápida identificação de ativos críticos, viabilizando respostas mais ágeis para evitar falhas inesperadas e reduzir impactos nos processos produtivos.

A utilização da clusterização na gestão de ativos e manutenção preditiva não apenas melhora a confiabilidade dos ativos industriais, mas também fortalece a gestão de riscos, pois possibilita a antecipação de falhas e a implementação de estratégias proativas para otimização dos ativos. Assim, essa técnica reforça a importância da Inteligência Artificial na gestão de ativos, tornando a operação industrial mais eficiente, resiliente e alinhada aos princípios da Indústria 4.0.

A gestão de ativos de engenharia é imprescindível para a administração eficiente de infraestruturas, equipamentos e instalações industriais, garantindo a otimização do ciclo de vida dos ativos e a redução de custos operacionais. Segundo Almeida (2023), esse campo combina práticas de engenharia, economia e gestão estratégica, permitindo que os ativos sejam monitorados, mantidos e substituídos de forma planejada e baseada em dados.

O autor enfatiza que, em um contexto onde os ativos envelhecem e se degradam naturalmente, a tomada de decisão baseada em dados é necessária para evitar falhas inesperadas e melhorar a eficiência operacional. Essa abordagem se alinha diretamente com o uso de Inteligência Artificial (IA) e clusterização, já que a análise preditiva permite identificar padrões de degradação e prever quando um equipamento precisará de manutenção ou substituição.

A digitalização e a automação, são pilares na evolução da gestão de ativos. Almeida (2023) ressalta que a integração de tecnologias como IoT, sensores inteligentes e Big Data possibilita um acompanhamento contínuo do desempenho dos ativos industriais, de forma que forneçam dados para otimizar manutenção, investimentos e riscos operacionais.

Outro ponto destacado pelo autor é a capacidade dos ativos industriais de se manterem operacionais e funcionais mesmo diante de situações adversas, como falhas mecânicas, desastres naturais ou ataques cibernéticos. O autor argumenta que um sistema de gestão de ativos, apoiado em dados e automação, pode aumentar a vida útil dos equipamentos e reduzir impactos operacionais, garantindo maior segurança e confiabilidade nos processos produtivos.

Por fim, Almeida (2023) reforça que a gestão de ativos não deve se limitar à manutenção corretiva ou preventiva, mas sim evoluir para uma abordagem integral, na qual a tecnologia e a

inteligência de dados são suportes na tomada de decisões estratégicas.

O estudo de Raptis, Passarella e Conti (2019) explora a gestão de dados na Indústria 4.0, destacando como a digitalização dos ativos industriais impulsiona a eficiência operacional e a tomada de decisões estratégicas. A pesquisa revisa o estado da arte em gestão de dados, integração de sistemas inteligentes e análise preditiva, além de abordar desafios e oportunidades na adoção de tecnologias como Inteligência Artificial (IA), Big Data e Internet das Coisas (IoT).

Os autores destacam que, na gestão de ativos industriais, os dados gerados por sensores e dispositivos conectados podem ser usados para melhorar o monitoramento, prever falhas e otimizar os ciclos de vida dos equipamentos. Além disso, o estudo discute estratégias de armazenamento, processamento e análise de dados para garantir segurança, interoperabilidade e escalabilidade nos sistemas de gestão de ativos.

No contexto da gestão de ativos, o estudo reforça a importância de estruturas eficientes para armazenamento e processamento de dados, permitindo que empresas possam correlacionar informações sobre o desempenho dos ativos, seu desgaste ao longo do tempo e sua criticidade operacional.

Além disso, os autores abordam a necessidade de padrões e protocolos para a gestão de dados industriais, garantindo que os sistemas conectados à Indústria 4.0 operem de forma integrada e segura. Na gestão de ativos, esse aspecto permite a criação de modelos mais precisos para análise de risco, previsão de falhas e otimização de investimentos em manutenção e substituição de equipamentos.

Por fim, o artigo conclui que, apesar dos avanços tecnológicos, ainda há desafios significativos no gerenciamento de dados industriais, sendo primordial o desenvolvimento de estratégias mais desenvolvidas de IA e aprendizado de máquina para transformar dados brutos em conhecimento acionável.

A tomada de decisão na gestão de ativos industriais têm evoluído para abordagens proativas e baseadas em dados, impulsionadas pela Inteligência Artificial (IA) e aprendizado de máquina. O estudo de Bousdekis *et al.* (2015) propõe um framework de decisão proativa voltado para a manutenção baseada em condição (CBM - Condition-Based Maintenance), permitindo que intervenções sejam realizadas antes que falhas comprometam a operação dos ativos.

O framework desenvolvido pelos autores é estruturado em dois espaços complementares: o espaço de informação, onde os dados operacionais dos ativos são coletados e processados, e o espaço de decisão, onde essas informações são utilizadas para gerar recomendações inteligentes para manutenção e gestão de riscos. A partir de sensores distribuídos em equipamentos, o sistema realiza análises preditivas e correlaciona padrões operacionais, otimizando a identificação de falhas iminentes e priorização de intervenções.

A aplicação desse modelo fortalece a gestão dos ativos, pois reduz paradas inesperadas e custos de manutenção, ao mesmo tempo em que melhora a disponibilidade e confiabilidade

operacional. Essa abordagem se alinha diretamente com a proposta do presente estudo, que busca utilizar clusterização e IA para correlacionar variáveis críticas, como ciclo de vida, criticidade dos ativos e custos operacionais, promovendo uma gestão inteligente e automatizada.

Além disso, Bousdekis *et al.* (2015) destacam a importância de estruturar modelos preditivos que não apenas identifiquem falhas, mas também forneçam suporte para decisões estratégicas. Isso reforça o papel da análise avançada de dados na otimização da gestão de ativos, garantindo que recursos sejam alocados de forma mais eficiente e que ativos críticos sejam monitorados continuamente para evitar falhas inesperadas.

Assim, o estudo demonstra a necessidade de evoluir de uma gestão reativa ou preventiva para uma abordagem inteligente e baseada em análise preditiva, demonstrando que a integração de IA e machine learning podem ser determinantes para aumentar a eficiência e competitividade das operações industriais na Indústria 4.0.

A integração de IA tem transformado a forma como os ativos industriais são geridos. Tecnologias ascendentes como o aprendizado de máquina e a clusterização permitem às empresas identificar padrões ocultos nos dados, otimizar o uso de seus ativos, reduzir custos operacionais e garantir a continuidade dos processos produtivos. A clusterização oferece formas para agrupar ativos com comportamentos similares, facilitando a priorização de intervenções e a análise preditiva de falhas. Apesar dos desafios inerentes à implementação dessas tecnologias avançadas, os benefícios a longo prazo, como o aumento da competitividade, a melhoria na tomada de decisões e a redução de falhas inesperadas, tornam essas inovações indispensáveis no cenário industrial atual.

### 3 METODOLOGIA

#### 3.1 Base de Dados

Para a coleta dos dados, a equipe de manutenção responsável pela planta realizou o preenchimento de duas planilhas no Excel: uma com os dados dos ativos de automação e outra com os equipamentos, que são compostos pelos ativos, com informações úteis para gestão de risco.

Sobre os dados dos ativos, estão dispostas informações como por exemplo: Aplicação, Família, Modelo, Descrição do item, Fabricante, Código do Ativo Instalado, Status do Ciclo de Vida do Ativo Instalado, Quantidade de Sobressalentes, Quantidade Instalada, Código do Sobressalente, Código do Sucessor, Valor do ativo, Valor Total do Ativo. Na Tabela 1 são exibidas os dados dos ativos como variáveis e a qual tipo pertencem originalmente.

Tabela 1 – Descrição das Variáveis Relacionadas à Gestão de Ativos.

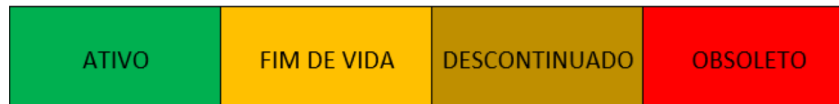
<b>Variável</b>	<b>Tipo</b>
Aplicação	Textual
Família	Categórica
Modelo	Textual
Descrição do item	Textual
Fabricante	Categórica
Código do Ativo Instalado	Textual
Status do Ciclo de Vida do Ativo Instalado	Categórica
Quantidade de Sobressalentes	Numérica
Quantidade Instalada	Numérica
Código do Sobressalente	Textual
Código do Sucessor	Textual
Valor do Ativo	Numérica
Valor Total do Ativo	Numérica

Cada uma dessas informações expõe características sobre o ativo ou papel que ele desempenha na planta. A aplicação indica onde o ativo está aplicado, como um CLP de um equipamento específico. A família é o grupo ao qual ele pertence, sendo componente de um CLP, um inversor, switches, entre outros. O modelo expõe tecnicamente o que é o ativo, como uma fonte de alimentação de um CLP ou um módulo de entrada digital. A descrição do item, por sua vez, trata com mais detalhes e informações sobre o ativo. O fabricante é o responsável pela fabricação do ativo. O código do ativo instalado é fornecido pela empresa fabricante, e por meio dele podemos consultar informações como status do ciclo de vida, código do sucessor (se houver) e valor do ativo.

Sobre as classificações de um ciclo de vida, em primeiro lugar existem os produtos “Ativos”, que são as versões mais recentes dos produtos, que não existem sucessores ou data de descontinuação anunciada, e há suporte total do fabricante. Os produtos em “Fim de Vida” tem

data de descontinuação anunciada, o suporte é total, mas versões mais recentes podem existir. Produtos “Descontinuados” são aqueles em que a data de descontinuação anunciada se passa, com isso, o produto deixa de ser manufaturado, porém com serviços de reparo ainda disponíveis. E por fim, produtos “Obsoletos” são aqueles que já ultrapassaram completamente sua vida útil e não possuem mais suporte técnico ou peças disponíveis no mercado. Na Figura 3 estão representadas as classificações dos possíveis ciclos de vida dos ativos.

Figura 3 – Ciclo de vida dos ativos.



Fonte: Elaborado pelo autor, 2025.

A quantidade de sobressalentes é o número de peças de reposição disponíveis na planta. A quantidade instalada é o número de ativos instalados e em operação na planta. O Código do sobressalente é referente a peça de reposição, assim como o código do sucessor é o código da peça mais atual disponível no mercado. Por fim, o valor do ativo é o preço do ativo mais recente disponível, assim como o valor total é o valor do ativo multiplicado pelo número de peças em operação na planta industrial.

A gestão de riscos traz informações fundamentais para a planta, como Equipamento, Criticidade, Probabilidade e Impacto. O equipamento é o local, como painel ou máquinas, onde os componentes estão em operação. A criticidade é um nível de medida da gravidade da parada do equipamento ou ativo para a planta. A probabilidade é uma escala percentual da ocorrência de uma falha do ativo ou equipamento. Por fim, o impacto é o impacto gerado pela falha do ativo ou equipamento para a planta industrial. Na Tabela 2 são exibidas as informações referentes à gestão de riscos e a qual tipo as variáveis pertencem.

Tabela 2 – Descrição das Variáveis Relacionadas à Gestão de Riscos.

Variável	Tipo
Equipamento	Textual
Criticidade	Categórica
Probabilidade	Categórica
Impacto	Categórica

## 3.2 Tratamento dos Dados

### 3.2.1 Organização

A organização da base de dados foram etapas fundamentais para o desenvolvimento deste trabalho, garantindo a qualidade das informações utilizadas na análise e no treinamento

do modelo (Valentim, 2001). A base de dados, obtida a partir da planilha de gestão de ativos e da planilha de gestão de riscos continham informações pilares para o trabalho, mas estavam separadas e necessitavam de integração, limpeza e padronização para serem utilizadas de forma eficiente.

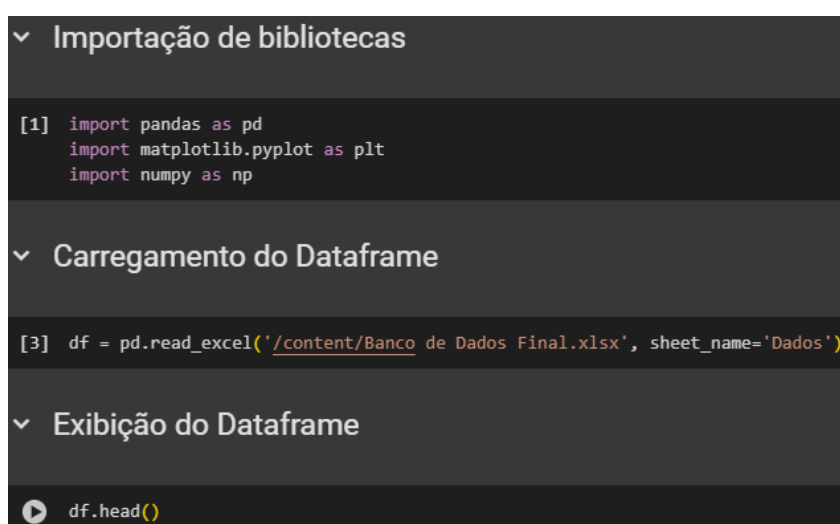
Inicialmente, as informações foram consolidadas em uma única base de dados, na qual as variáveis foram integradas. Durante esse processo, foi realizada a atualização dos valores referentes ao status do ciclo de vida dos ativos, para refletir a situação mais recente dos equipamentos. Adicionalmente, os dados relacionados aos valores dos ativos também foram atualizados para garantir maior precisão nas análises. Esse processo resultou em uma base de dados (composta por 678 dados diferentes) organizada e adaptada para os próximos passos do trabalho.

### 3.2.2 *Tratamento*

A fim de desenvolver uma melhor análise de dados, foi utilizada a linguagem de programação Python (McKinney, 2012). Para facilitar e otimizar o processo, foi utilizado o interpretador Jupyter Notebook, do ambiente virtual Google Colab.

Dentro do Jupyter Notebook foi possível desenvolver de forma mais organizada e estruturada, pois dentro da ferramenta é possível executar diferentes blocos de códigos na ordem escolhida. Dessa forma, tornou-se o mais adequado para tratamento de dados, visualizações gráficas e análises. Na Figura 4 está representado um trecho do código desenvolvido no trabalho.

Figura 4 – Trecho do código em Jupyter Notebook.



```

▼ Importação de bibliotecas

[1] import pandas as pd
    import matplotlib.pyplot as plt
    import numpy as np

▼ Carregamento do Dataframe

[3] df = pd.read_excel('/content/Banco de Dados Final.xlsx', sheet_name='Dados')

▼ Exibição do Dataframe

df.head()
```

Fonte: Elaborado pelo autor, 2025.

Nessa linguagem, são oferecidos diversos suportes pela utilização de bibliotecas, que contribuem significativamente os processos de criação dos algoritmos (Stančin; Jović, 2019). A biblioteca Python Pandas foi utilizada a fim de importar os arquivos, realizar checagens e

manipular dados. Essa biblioteca foi essencial para verificação de dados nulos ou faltantes e garantia da qualidade da base de dados.

Como a base de dados incluía variáveis categóricas e numéricas, foi necessário mapear as variáveis categóricas para transformar em numéricas, com a finalidade de serem utilizadas nos cálculos e análises (Santos *et al.*, 2024). Essa etapa foi essencial, já que os métodos de clusterização, como o K-Means, requerem variáveis numéricas para realizar cálculos de forma adequada. Na Tabela 3 está representado como foi realizado o mapeamento das variáveis categóricas para numéricas.

Tabela 3 – Mapeamento das Variáveis Categóricas para Numéricas.

Variável	Valor Categórico	Valor Numérico
Status do Ciclo de Vida do Ativo Instalado	ATIVO	1
	FIM DE VIDA	2
	DESCONTINUADO	3
	OBSOLETO	4
Criticidade	1	4
	A	3
	B	2
	C	1
Probabilidade	1 - Raro	1
	2 - Não impossível	2
	3 - Possível	3
Impacto	1 - Insignificante	1
	2 - Baixo	2
	3 - Moderado	3
	4 - Elevado	4
	5 - Crítico	5

Além disso, foi utilizada a biblioteca *MinMaxScaler*, com o objetivo de normalizar os dados presentes em cada coluna da base de dados. Essa biblioteca aplica cálculos de normalização que ajustam os valores de forma proporcional entre 0 e 1. Esse processo evita que diferenças de escala entre os atributos torne o modelo tendencioso (Santana; Pontes, 2020). Apesar da normalização alterar os valores individuais dos dados, ela mantém os parâmetros estatísticos gerais das colunas, preservando a relação entre as variáveis e assegurando a integridade da análise.

### 3.3 Definição e Cálculo do Critério de Prioridade

Após o tratamento dos dados, foi desenvolvido um critério denominado Prioridade, que para o seu cálculo, considera múltiplos fatores relacionados à gestão de risco e gestão dos ativos. Esse cálculo integra variáveis definidas como pilares à gestão de ativos, ponderadas por pesos que apontam sua importância relativa, expresso pela Equação 3.1.

$$\text{Prioridade} = w_1 \times \text{Risco\_Total} + w_2 \times \text{Criticidade} + w_3 \times \text{Ciclo\_de\_Vida} + w_4 \times \text{Relacao\_Instalada\_Sobressalente} \quad (3.1)$$

Dentre os componentes do cálculo da prioridade, temos as seguintes descrições: O Risco Total é obtido pelo produto entre o impacto e a probabilidade. A Criticidade e o Status do Ciclo de Vida do Ativo são representados pelos valores mapeados numericamente e normalizados anteriormente. A Relação Instalada/Sobressalente é uma variável que avalia a relação entre a quantidade instalada do ativo e a disponibilidade de sobressalentes. Essa variável é representada pela Equação 3.2.

$$\text{Rel}_{\text{IS}} = \frac{\text{Quantidade Instalada}}{\text{Quantidade de Sobressalentes} + \varepsilon}, \quad \text{com } \varepsilon > 0 \quad (3.2)$$

Para encontrar os valores dos pesos que compõem o cálculo da prioridade, foi desenvolvido um código que encontra os melhores pesos ( $w_1, w_2, w_3, w_4$ ) a partir de iterações, garantindo que esses valores sejam ajustados de maneira otimizada utilizando o *silhouette score*, uma métrica que avalia a qualidade da clusterização (Lopes; Cavique, 2022). Essa métrica mede a coesão interna dos clusters e a separação entre eles, sendo definida pela Equação 3.3.

$$S = \frac{b - a}{\max(a, b)} \quad (3.3)$$

Onde:

- a = distância média entre um ponto e os outros pontos do cluster pertencente (coesão)
- b = distância média entre um ponto e o centroide do cluster mais próximo (separação).

O valor encontrado pode variar de -1 a 1, sendo que 1 indica clusters bem separados. 0 indica sobreposição entre clusters e -1 indica elementos alocados em clusters errados.

Com isso, o processo de otimização dos pesos começa com a definição de um intervalo de possíveis valores para os pesos. O processo de otimização do peso é sujeito às seguintes restrições: A soma dos pesos estar dentro da escala de normalização das variáveis utilizadas no cálculo da prioridade e respeitar a hierarquia lógica, onde um peso possui maior relevância que o outro, sendo  $w_4 > w_1 > w_2 > w_3$ .

Para cada combinação de pesos válida (região viável), o código calcula a prioridade dos ativos e aplica o algoritmo de clusterização K-means. Logo após, é calculado o *silhouette score*. A função da otimização dos pesos é maximizar o *silhouette score*.

Ao final do processo, os melhores pesos encontrados são utilizados para recalcular a prioridade dos ativos de forma definitiva, garantindo que os agrupamentos resultantes sejam os

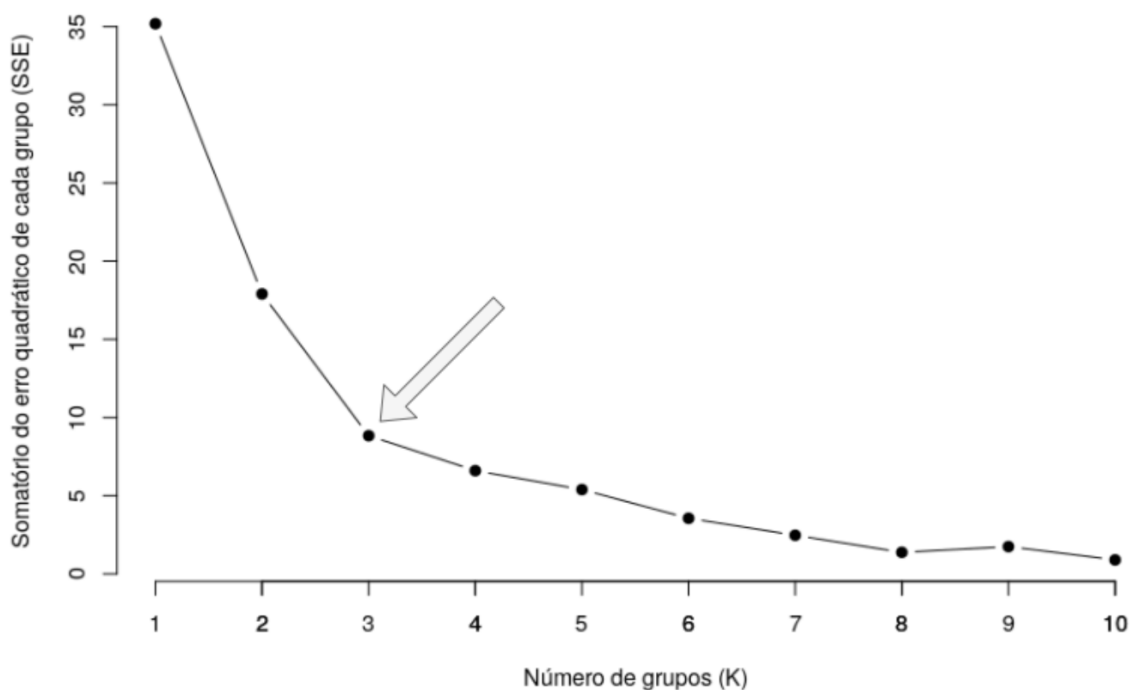
mais coerentes possíveis. Esse método permite encontrar pesos ideais de forma automatizada, garantindo que a classificação dos ativos siga critérios bem definidos.

### 3.4 Clusterização

Para segmentar e os ativos de automação baseado em suas características, foi utilizada a clusterização, técnica de aprendizado não supervisionado. Esse método identifica padrões nos dados e auxilia na tomada de decisão sobre a priorização de manutenção e substituição dos ativos (Costa *et al.*, 2024). Além disso, após a clusterização pode-se visualizar os agrupamentos gerados para melhor interpretação dos resultados.

Para aplicar a clusterização, é necessário em primeiro lugar, definir a quantidade ideal de clusters nos quais os ativos serão divididos. Para isso, foi utilizado o Método do Cotovelo (Elbow-Method), que é uma forma de descobrir quantos grupos (clusters) devemos utilizar ao usar o K-Means. O método consiste em rodar o algoritmo para diferentes valores de K e calcular a soma dos erros quadrados dentro dos clusters (WSS, do inglês Within-Cluster Sum of Squares). Quando você plota o WSS em função do número de clusters, a forma do gráfico geralmente se assemelha a um cotovelo. O ponto onde a redução no WSS começa a diminuir drasticamente é considerado o número ideal de clusters. Esse critério evita a criação de clusters excessivos ou insuficientes, garantindo uma segmentação eficiente dos ativos (Bholowalia; Kumar, 2014). Na Figura 5 está um exemplo do uso desse método.

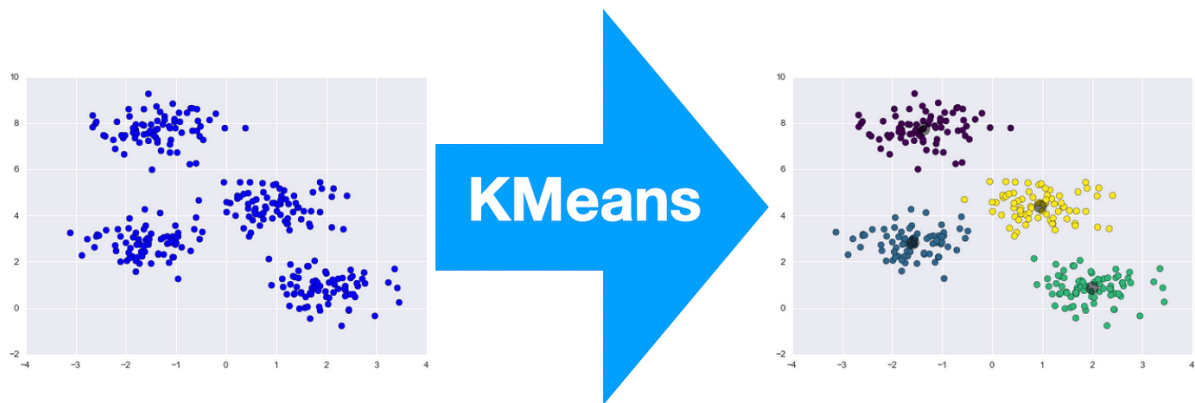
Figura 5 – Exemplificação do Método do Cotovelo.



Com o valor de K definido, foi aplicado o algoritmo K-Means, um dos métodos mais utilizados para agrupamento de dados (Ikotun *et al.*, 2023). Seu objetivo é minimizar a variabilidade dentro de cada grupo e maximizar a separação entre os grupos, garantindo que os pontos dentro de um mesmo cluster sejam mais parecidos entre si do que com os de outros clusters.

O algoritmo K-Means inicia escolhendo K pontos iniciais, centroides. Em seguida, cada ponto do conjunto de dados é atribuído ao centroide mais próximo, geralmente com base na distância euclidiana. Após essa atribuição inicial, os centroides são recalculados como a média dos pontos de cada cluster. Esse processo de atribuição e reajuste dos centroides se repete iterativamente até que os centroides não mudem mais significativamente ou até atingir um número máximo de iterações, garantindo que os grupos estejam bem definidos e estabilizados. A Figura 6 demonstra o funcionamento do K-Means.

Figura 6 – Exemplificação do Método K-Means.



Fonte: Barros, 2021.

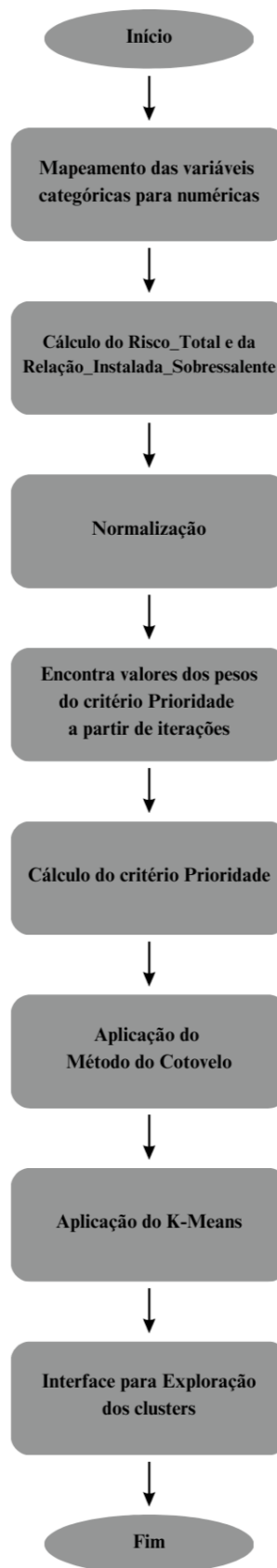
O algoritmo K-Means foi escolhido para este trabalho devido à sua eficiência computacional e ampla aplicabilidade em problemas de segmentação, especialmente em cenários industriais onde a categorização de ativos pode otimizar processos de manutenção e substituição (Ikotun *et al.*, 2023). Além disso, o K-Means é um dos métodos mais utilizados para agrupamento de dados devido à sua simplicidade e capacidade de encontrar padrões em grandes volumes de informação (Bholowalia; Kumar, 2014). Sua natureza iterativa permite a minimização da variabilidade dentro de cada grupo e a maximização da separação entre os clusters. Desta forma, garante que os ativos classificados dentro de um mesmo grupo apresentem características semelhantes, facilitando a análise e tomada de decisões estratégicas (Costa *et al.*, 2024).

### 3.5 Interface para Exploração dos Clusters

Para facilitar a análise e aplicação prática dos resultados da clusterização, foram implementadas duas abordagens principais: a exportação dos dados para um arquivo Excel e a criação de uma interface interativa. Os dados dos clusters, identificados pelo K-Means, foram exportados

para um arquivo Excel, com cada cluster salvo em uma aba separada, permitindo que os gestores acessem e manipulem as informações facilmente. Além disso, foi desenvolvida uma ferramenta interativa utilizando a biblioteca ipywidgets, que permite aos usuários explorar os clusters, selecionando diferentes atributos e clusters para análise. Assim, o fluxograma do funcionamento do algoritmo está representado pela Figura 7.

Figura 7 – Fluxograma do funcionamento do algoritmo.



## 4 RESULTADOS

### 4.1 Caracterização dos Dados e Pré-Processamento

A base de dados utilizada neste estudo contém informações detalhadas sobre ativos industriais de tecnologia da automação. Cada registro representa um ativo específico e suas características operacionais, financeiras e de riscos. As principais variáveis analisadas estão dispostas na Tabela 4:

Tabela 4 – Principais variáveis analisadas.

Variável	Descrição
Código do ativo	Identificação única do equipamento.
Modelo e fabricante	Características específicas do ativo.
Quantidade instalada	Número de unidades em operação.
Quantidade de sobressalentes	Número de peças disponíveis para reposição.
Valor do ativo	Custo individual de cada tipo de ativo.
Valor total do ativo	Custo agregado de cada tipo de ativo.
Criticidade do ativo	Classificação da importância do ativo no processo industrial.
Status do ciclo de vida do ativo	Indica se a classificação do ciclo de vida do ativo.
Impacto e probabilidade de falha	Parâmetros utilizados para calcular o risco total do ativo.
Relação instalada/sobressalente	Métrica que auxilia na gestão do estoque de peças de reposição.

Para permitir que os algoritmos de IA e aprendizado de máquina analisassem os dados corretamente, algumas variáveis foram convertidas para valores numéricos. Esse processo é necessário para que o modelo de clusterização K-Means possa identificar padrões nas variáveis. O mapeamento seguiu a estrutura apresentada na Tabela 3. Essa conversão foi realizada a fim de garantir que os modelos analíticos pudessem extrair por completo os dados disponíveis e fosse capaz de comparar os diferentes ativos.

A gestão de ativos requer indicadores quantitativos que auxiliem na tomada de decisões. Desta forma, dois índices foram calculados para avaliar os ativos: “Risco Total” e “Relação Instalada/Sobressalente”. O Risco Total de um ativo é calculado com base no produto das variáveis “Impacto” e “Probabilidade”. O “Impacto” corresponde a severidade do impacto que uma falha do ativo pode causar no processo industrial e a “Probabilidade” mede a chance de que o ativo falhe na operação. Desta forma, quanto maior o valor do Risco Total, maior a urgência de estratégias de mitigação. A equação adotada para o Risco Total está representada por 4.1.

$$\text{Risco Total} = \text{Impacto} \times \text{Probabilidade} \quad (4.1)$$

Outro indicador importante desenvolvido para a análise de ativos industriais é a Relação Instalada/Sobressalente, que mede a disponibilidade de peças sobressalentes em relação ao número de unidades instaladas. O cálculo dessa variável está representada pela Equação 3.2. As-

sim, é possível identificar a situação do estoque na planta industrial, o que impacta diretamente na gestão do estoque.

As variáveis presentes na base de dados possuem escalas distintas, o que pode afetar o modelo desenvolvido pelos algoritmos de clusterização. Por exemplo, o Status do Ciclo de Vida do Ativo pode variar de 1 a 4, conforme mapeamento, enquanto o Risco Total pode variar de 1 a 10. A Tabela 5 representa as estatísticas dos dados antes da normalização.

Tabela 5 – Estatísticas descritivas antes da normalização.

Estatística	Status do Ciclo de Vida do Ativo	Criticidade	Probabilidade	Impacto	Risco Total	Relação Instalada/Sobressalente
Contagem	678	678	678	678	678	678
Média	2,754	2,487	1,835	2,254	4,118	0,806
Desvio Padrão	1,099	0,661	0,903	0,893	2,318	0,773
Mínimo	1,000	1,000	1,000	1,000	1,000	0,500
1º Quartil	2,000	2,000	1,000	2,000	2,000	0,500
Mediana	3,000	2,000	1,000	2,000	4,000	0,667
3º Quartil	4,000	3,000	3,000	2,000	6,000	0,889
Máximo	4,000	4,000	3,000	5,000	10,000	13,000

Para corrigir esse problema e garantir que todas as variáveis tivessem a mesma escala na análise, foi aplicada a técnica de Min-Max Scaling, que normaliza os valores para um intervalo entre 0 e 1, conforme a Equação 4.2.

$$X_{\text{normalizado}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (4.2)$$

Isso garante que nenhuma variável faça com que a análise seja tendenciosa (Santana; Pontes, 2020), permitindo que a clusterização seja realizada de forma equilibrada. A Tabela 6 representa as estatísticas dos dados após a normalização.

Tabela 6 – Estatísticas descritivas após a normalização.

Estatística	Status do Ciclo de Vida do Ativo	Criticidade	Probabilidade	Impacto	Risco Total	Relação Instalada/Sobressalente
Contagem	678	678	678	678	678	678
Média	0,585	0,496	0,417	0,313	0,346	0,024
Desvio Padrão	0,366	0,220	0,451	0,223	0,258	0,062
Mínimo	0,000	0,000	0,000	0,000	0,000	0,000
1º Quartil	0,333	0,333	0,000	0,250	0,111	0,000
Mediana	0,667	0,333	0,000	0,250	0,333	0,013
3º Quartil	1,000	0,667	1,000	0,250	0,556	0,031
Máximo	1,000	1,000	1,000	1,000	1,000	1,000

O pré-processamento dos dados foi uma etapa realizada a fim de garantir a qualidade da análise e a eficácia dos algoritmos de IA. A normalização das variáveis possibilitou uma comparação diante de uma mesma escala e a transformação das variáveis para numéricas permitiu a inclusão das mesmas nos passos seguintes do trabalho. Com os dados devidamente estruturados, foi possível avançar para a etapa de cálculo do critério de prioridade.

## 4.2 Cálculo do Critério de Prioridade dos Ativos

A gestão de ativos industriais requer um critério que permita identificar quais ativos demandam intervenções mais urgentes. Para isso, foi desenvolvido um critério de priorização que combina múltiplos fatores do ativo, ponderados de forma otimizada para garantir melhor classificação em relação à prioridade dos ativos. Esse critério “Prioridade” é baseado em uma equação ponderada e otimizada por meio de um algoritmo de aprendizado de máquina, que ajusta os pesos das variáveis utilizando a métrica silhouette score para garantir a melhor clusterização dos ativos.

O critério “Prioridade” foi determinado a partir da Equação 3.1. Com isso, foi possível criar um indicador quantitativo que permite diferenciar ativos com base em sua necessidade de gestão, além de balancear fatores de risco, criticidade e disponibilidade de sobressalentes de forma otimizada. Assim, conseguindo agrupar os ativos em clusters bem definidos, facilitando a tomada de decisão.

Dado que diferentes combinações de pesos poderiam impactar a separação dos ativos, foi implementado um algoritmo de otimização que ajusta os pesos  $w_1, w_2, w_3, w_4$  a partir de iterações com o intuito de maximizar a qualidade da clusterização.

O algoritmo testa combinações de pesos no intervalo de 0,01 a 1,0 com passo de 0,01 garantindo que a soma total dos pesos seja igual a 1 para manter a escala das variáveis normalizadas e que os pesos sigam a hierarquia  $w_4 > w_1 > w_2 > w_3$ , que foi definida em conhecimentos prévios juntamente com a equipe responsável pela gestão da planta industrial. Com isso, o algoritmo faz o cálculo da Prioridade para cada conjunto de pesos usando a equação definida e aplica o K-Means utilizando o silhouette score que quanto maior o valor, melhor a separação entre clusters. A iteração continua até encontrar a melhor combinação de valores dos pesos para o maior valor possível do silhouette score que atenda às restrições dos pesos.

Finalizado as iterações do algoritmo de otimização, os pesos encontrados foram:

- $w_1=0,26$  (Risco Total) - Foi o fator mais relevante após a relação de sobressalentes, pois o risco impacta diretamente na operação;
- $w_2=0,25$  (Criticidade) - A criticidade teve um peso intermediário, influenciando na priorização, mas sem dominar o critério;
- $w_3=0,22$  (Status do Ciclo de Vida do ativo) - Teve peso muito baixo, pois o ciclo de vida dos ativos não são os mais críticos para manutenção imediata;
- $w_4 =0,27$  (Relação Instalada/Sobressalente) - Foi o fator mais relevante, pois a falta de sobressalentes impacta diretamente na capacidade de resposta a falhas.

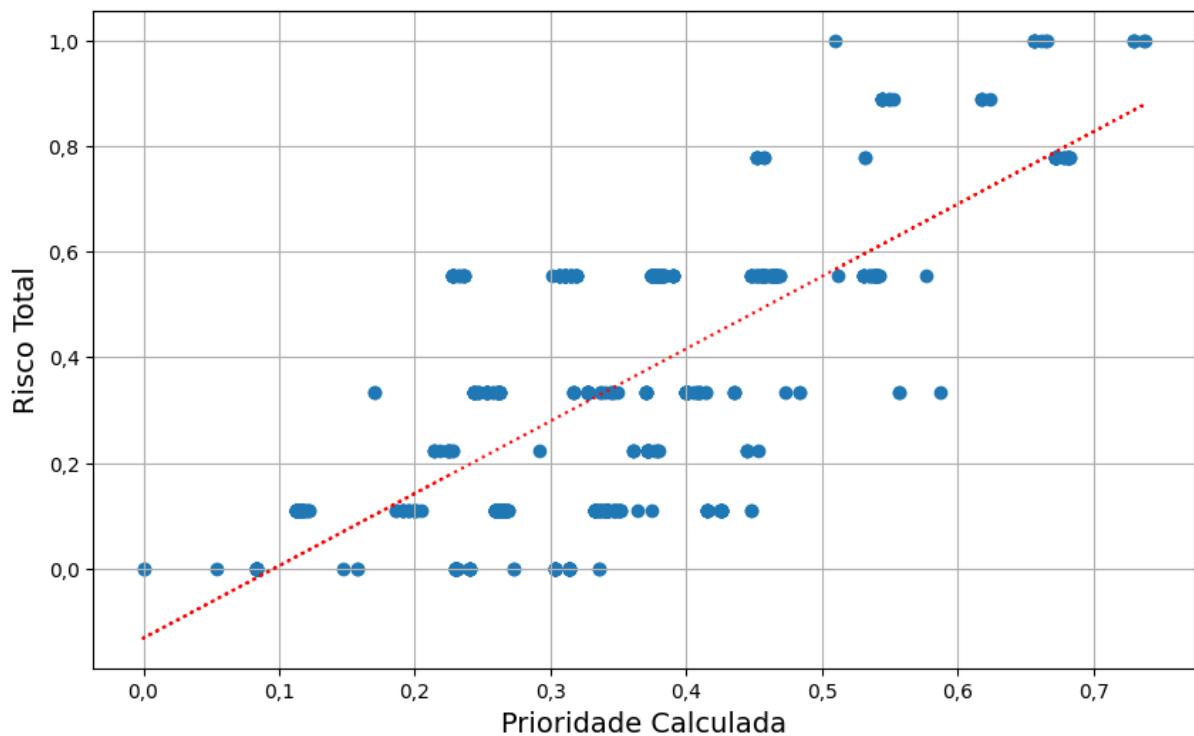
Desta forma, com os valores dos pesos encontrados, o critério da Prioridade é definido pela Equação 4.3.

$$\text{Prioridade} = 0,26 \times \text{Risco\_Total} + 0,25 \times \text{Criticidade} + 0,22 \times \text{Ciclo\_de\_Vida} + 0,27 \times \text{Relacao\_Instalada\_Sobressalente} \quad (4.3)$$

Com isso, o melhor silhouette score que atenda às restrições dos pesos foi 0,413. A clusterização foi ajustada para garantir que os ativos mais prioritários fossem corretamente identificados. A melhor separação entre a prioridade dos ativos foi obtida com os pesos  $\mathbf{w} = [0,26; 0,25; 0,22; 0,27]$ .

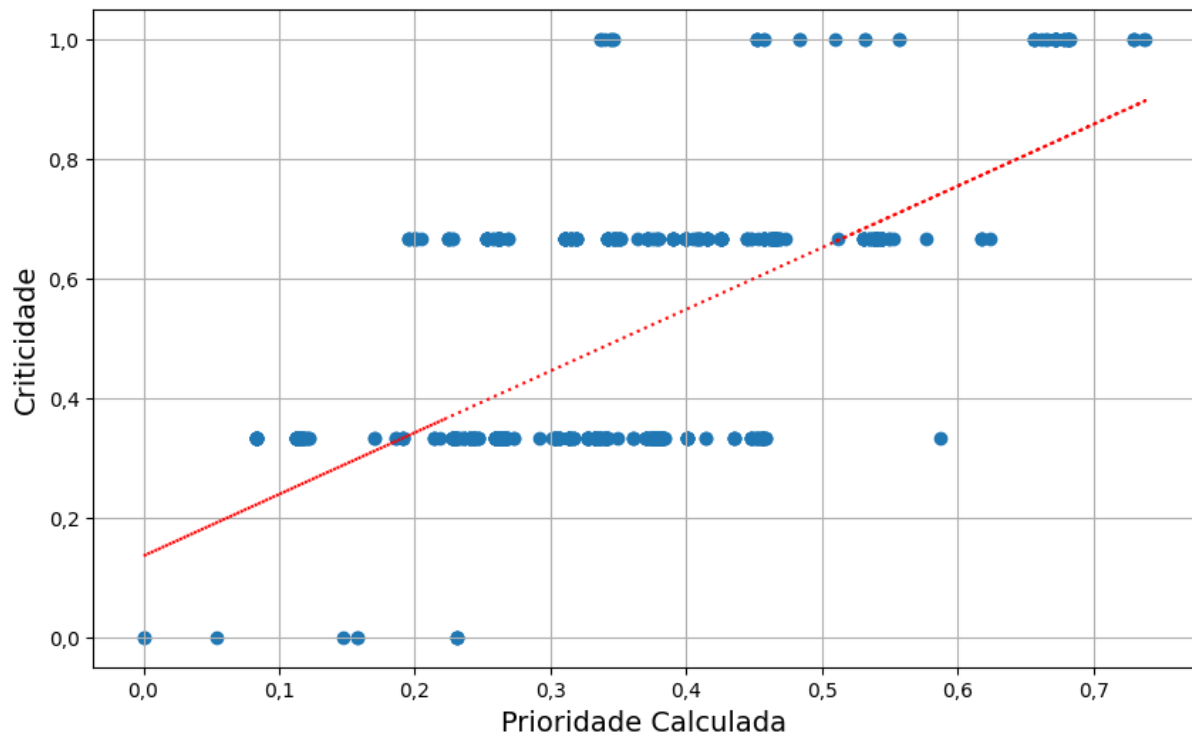
Para validar o impacto dos pesos na priorização dos ativos, foram gerados gráficos de dispersão com linha de tendência mostrando a relação entre prioridade calculada com risco total (Figura 8), assim como prioridade calculada e criticidade (Figura 9).

Figura 8 – Relação entre Prioridade e Risco Total.



Fonte: Elaborado pelo autor, 2025.

Figura 9 – Relação entre Prioridade e Criticidade.



Fonte: Elaborado pelo autor, 2025.

Após análise estatística, a relação entre o risco total e a prioridade apresentou um p-valor de  $2,80 \times 10^{-108}$ , enquanto para a relação entre a criticidade e a prioridade foi encontrado um p-valor de  $8,82 \times 10^{-77}$ . Desta forma, sendo valores menores que 0,05, é evidenciado estatisticamente que o critério de priorização dos ativos está significativamente associado a essas variáveis. Isso demonstra que aos ativos com maior risco e maior criticidade, foram atribuídos predominantemente valores mais altos de prioridade, enquanto aos ativos com menor criticidade e risco reduzido, foram atribuídos valores menores de prioridade. Logo, o critério de priorização reflete a realidade operacional dos ativos, assegurando que as intervenções de manutenção sejam planejadas de forma estratégica.

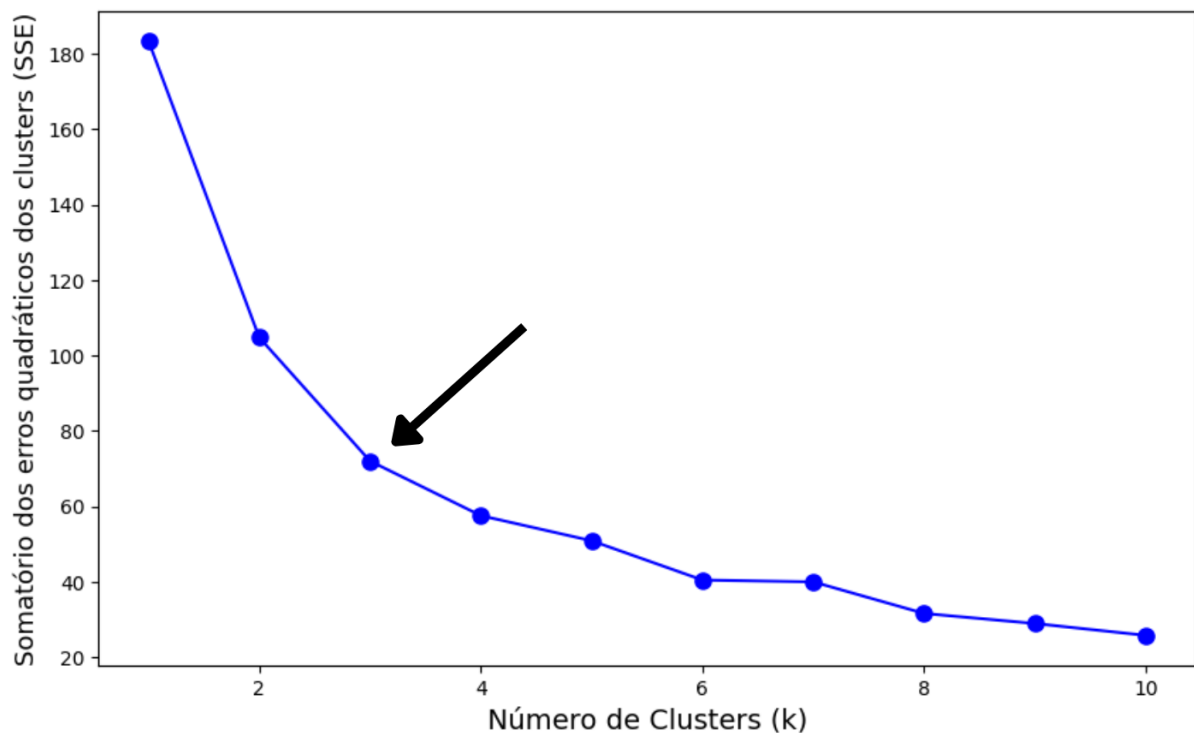
### 4.3 Identificação dos Clusters de Ativos de Automação

A clusterização permite segmentar os ativos em grupos distintos com características semelhantes. Para essa etapa foi utilizada a técnica de clusterização K-Means, que agrupou os ativos com base nas variáveis fornecidas, como risco total, criticidade, status do ciclo de vida do ativo, relação instalada/sobressalente e prioridade.

A definição do número ideal de clusters foi realizada por meio do método do cotovelo, buscando uma segmentação equilibrada e representativa dos ativos. Em seguida, por meio de características comuns o algoritmo K-Means realiza a clusterização, possibilitando análises para a gestão de ativos na automação industrial.

O método do cotovelo analisa a variação do somatório dos erros quadráticos dos clusters conforme o número K de clusters aumenta. Para isso, o K-Means foi aplicado com a faixa de K de 1 a 10 e com isso foi gerado um gráfico do método do cotovelo, analisando o ponto onde a redução do SSE diminui significativamente, ou seja, começa a estabilizar. O gráfico gerado (Figura 10) indicou que o número de clusters ideal foi 3, pois a partir desse ponto, a queda na redução do SSE torna-se menos significativa. Assim, os ativos se agrupam em três clusters distintos, otimizando a separação sem a criação de clusters artificiais.

Figura 10 – Gráfico do Método do Cotovelo.



Fonte: Elaborado pelo autor, 2025.

#### 4.4 Caracterização dos Clusters

Após a definição do número de clusters, foi implementado o algoritmo K-Means, um dos mais utilizados como técnica de aprendizado não supervisionado em agrupamento de dados. A implementação utilizou os parâmetros otimizados a partir do maior silhouette score possível. Os três clusters foram criados com as seguintes variáveis:

- Status do Ciclo de Vida do Ativo Instalado;
- Criticidade;
- Risco\_Total;
- Relacao\_Instalada\_Sobressalente;

- Prioridade;

Inicialmente o algoritmo escolhe aleatoriamente K pontos como centros iniciais dos clusters. Cada ativo é alocado ao cluster cujo centroide está mais próximo. Após isso, o K-Means faz uma recalibração, recalculando os centroides com base nos novos agrupamentos, repetindo esse processo até que os clusters fiquem estabilizados. Desta forma os ativos semelhantes são agrupados.

Após isso, foram formados os três clusters, em que a ordem e numeração dos mesmos foi determinada pelo próprio algoritmo, sendo eles: Cluster 0, Cluster 1 e Cluster 2. Assim, foi possível caracterizar os clusters gerados. A distribuição do número de ativos por cluster está representada na Figura 11. Com relação à prioridade, o Cluster 0 representa os ativos de baixa prioridade, o Cluster 1 média prioridade e o Cluster 2 alta prioridade.

Figura 11 – Distribuição de Ativos por Cluster.



Fonte: Elaborado pelo autor, 2025.

No primeiro cluster gerado, Cluster 0, dentre as estatísticas exibidas pela Tabela 7, é indicado ativos com ciclos de vida prioritários e com menor disponibilidade de sobressalentes, com maior flexibilidade de gestão e menor prioridade. Sua prioridade média é 0,228 e relação instalada/sobressalente média de 0,011.

Tabela 7 – Estatísticas descritivas dos ativos do Cluster 0.

Estatística	Valor do Ativo	Valor Total do Ativo	Status do Ciclo de Vida	Criticidade	Risco Total	Relação Instalada/Sobressalente	Prioridade
Contagem	195	195	195	195	195	195	195
Média	6.335,90	12.742,87	0,063	0,506	0,325	0,011	0,228
Desvio Padrão	10.393,96	26.068,11	0,131	0,218	0,222	0,018	0,098
Mínimo	626,00	626,00	0	0	0	0	0
1º Quartil	1.864,59	3.749,40	0	0,333	0,111	0	0,120
Mediana	3.151,20	6.160,00	0	0,333	0,333	0	0,236
3º Quartil	6.321,00	10.559,20	0	0,667	0,556	0,020	0,309
Máximo	77.662,26	232.986,78	0,333	1	1	0,200	0,510

O Cluster 1 representa os ativos de prioridade intermediária. Sua prioridade média é de 0,347 e a variável do status do ciclo de vida corresponde a 0,785. As estatísticas do Cluster 1 estão representadas pela Tabela 8.

Tabela 8 – Estatísticas descritivas dos ativos do Cluster 1.

Estatística	Valor do Ativo	Valor Total do Ativo	Status do Ciclo de Vida	Criticidade	Risco Total	Relação Instalada/Sobressalente	Prioridade
Contagem	365	365	365	365	365	365	365
Média	4.627,64	8.894,24	0,785	0,399	0,251	0,032	0,347
Desvio Padrão	9.568,24	21.017,34	0,160	0,148	0,207	0,078	0,066
Mínimo	95,00	95,00	0,667	0	0	0	0,147
1º Quartil	665,83	1.958,47	0,667	0,333	0,111	0	0,314
Mediana	2.143,00	3.618,18	0,667	0,333	0,111	0,020	0,349
3º Quartil	4.731,02	8.866,50	1	0,333	0,556	0,037	0,382
Máximo	144.549,15	224.000,00	1	0,667	0,556	1	0,587

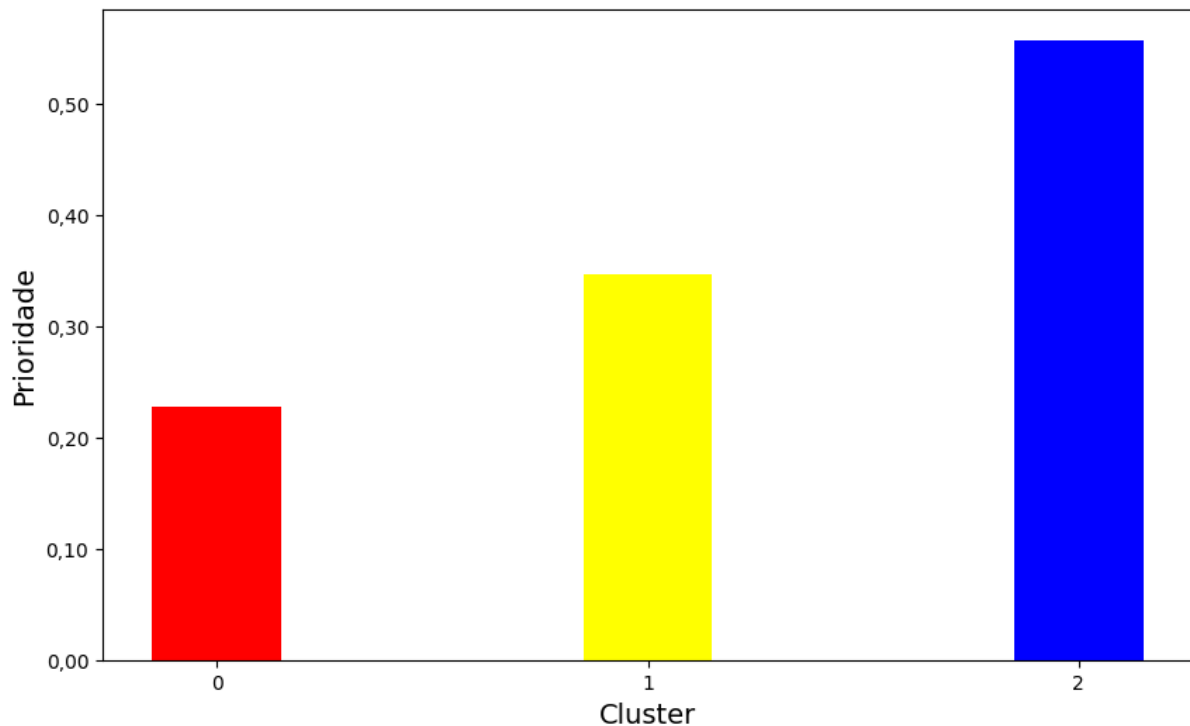
O Cluster 2 representa os ativos de maior criticidade e risco total reduzido, com maior risco operacional. Sua prioridade média é de 0,558 e risco total médio 0,676. As estatísticas do Cluster 2 estão representadas pela Tabela 9.

Tabela 9 – Estatísticas descritivas dos ativos do Cluster 2.

Estatística	Valor do Ativo	Valor Total do Ativo	Status do Ciclo de Vida	Criticidade	Risco Total	Relação Instalada/Sobressalente	Prioridade
Contagem	118	118	118	118	118	118	118
Média	5.039,02	14.846,99	0,825	0,777	0,676	0,023	0,558
Desvio Padrão	13.024,78	50.784,39	0,178	0,157	0,174	0,045	0,087
Mínimo	95,00	95,00	0,333	0,667	0,333	0	0,458
1º Quartil	323,75	1.385,01	0,667	0,667	0,556	0	0,466
Mediana	1.574,35	4.383,80	0,667	0,667	0,556	0,020	0,540
3º Quartil	4.383,80	9.075,02	1	1	0,778	0,030	0,657
Máximo	110.000,00	440.000,00	1	1	1	0,440	0,738

A divisão dos ativos em três clusters foi eficaz, já que foi validada pelo método do coto-velo e maior *silhouette score* possível dentro das condições estabelecidas, permitindo distinção entre prioridade baixa, média e alta (Figura 12). A clusterização demonstrou ser uma ferramenta capaz de auxiliar a gestão de ativos, incluindo fatores operacionais e financeiros. Assim, a integração da IA nesse contexto traz maior confiabilidade e desempenho.

Figura 12 – Prioridade média em relação a cada cluster.



Fonte: Elaborado pelo autor, 2025.

#### 4.5 Análise dos Clusters e Relevância para a Gestão de Ativos

Para analisar os clusters, foram gerados gráficos individuais explorando a distribuição das principais variáveis e o impacto das mesmas com relação a gestão de ativos de automação. A Figura 13 representa a distribuição dos ativos por status do ciclo de vida dentro de cada cluster.

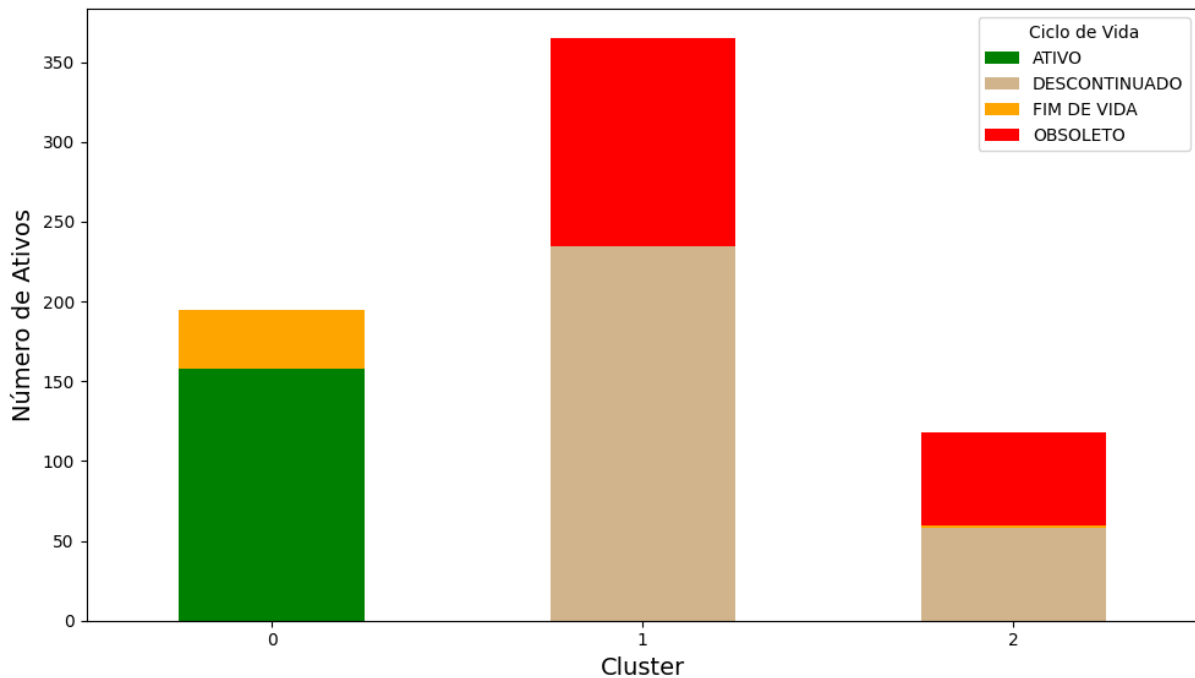
Observa-se que o Cluster 0 concentra a maior parte dos ativos classificados como "Ativo", indicando que esse grupo contém equipamentos mais recentes, que ainda possuem suporte completo do fabricante. Esse cenário sugere menor necessidade de substituições imediatas e menor risco de indisponibilidade de suporte técnico.

Por outro lado, os Clusters 1 e 2 apresentam uma proporção significativa de ativos classificados como "Descontinuado" e "Obsoleto", evidenciando que esses grupos possuem equipamentos mais antigos, que podem estar fora de linha e sem suporte do fabricante. No Cluster 1, há uma predominância de ativos "Descontinuados", enquanto no Cluster 2, observa-se uma quantidade expressiva de ativos "Obsoletos", reforçando a necessidade de ações estratégicas para substituição ou gestão especializada desses equipamentos.

Diante desse cenário, a predominância de ativos "Descontinuados" e "Obsoletos" nos Clusters 1 e 2 destaca a importância de um planejamento antecipado para substituições, a fim de minimizar impactos operacionais decorrentes da indisponibilidade de peças ou suporte técnico. Dessa forma, a análise dos clusters permite direcionar estratégias de manutenção e renovação

dos ativos, garantindo maior confiabilidade e continuidade operacional.

Figura 13 – Distribuição do Ciclo de Vida dos Ativos em cada Cluster.



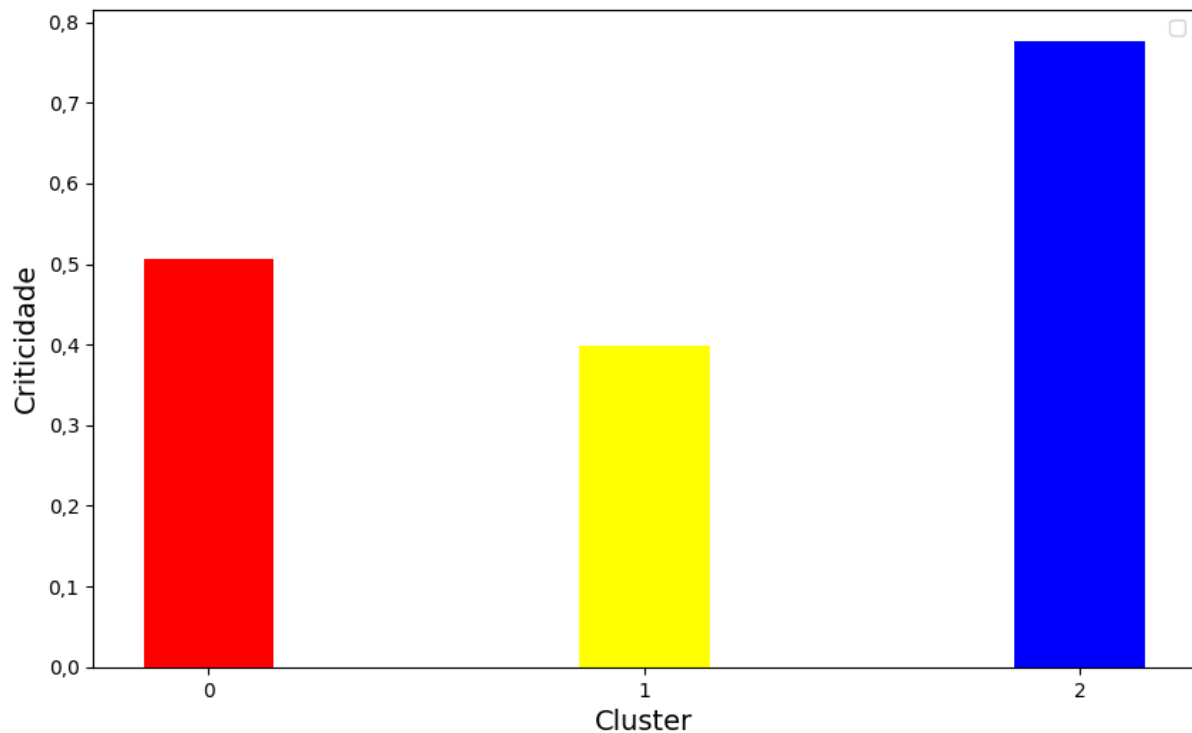
Fonte: Elaborado pelo autor, 2025.

A Figura 14 apresenta a criticidade média dos ativos dentro de cada cluster. Observa-se que o Cluster 2 possui a maior criticidade média entre os grupos, indicando que os ativos desse cluster têm um impacto operacional significativo e devem receber atenção especial na gestão de manutenção e mitigação de riscos.

O Cluster 0 apresenta uma criticidade intermediária, sugerindo que seus ativos possuem importância moderada dentro do contexto operacional. Já o Cluster 1 possui a menor criticidade média, o que indica que os ativos desse grupo possuem menor impacto na operação e podem não demandar ações emergenciais em termos de substituição ou manutenção.

Dessa forma, a análise reforça a necessidade de priorização dos ativos do Cluster 2 em estratégias de manutenção preventiva e corretiva, de modo a evitar falhas que possam comprometer a operação. Além disso, a criticidade intermediária do Cluster 0 sugere que seus ativos devem ser monitorados de forma estratégica, enquanto os ativos do Cluster 1 podem demandar um planejamento menos intensivo.

Figura 14 – Criticidade Média por Cluster.



Fonte: Elaborado pelo autor, 2025.

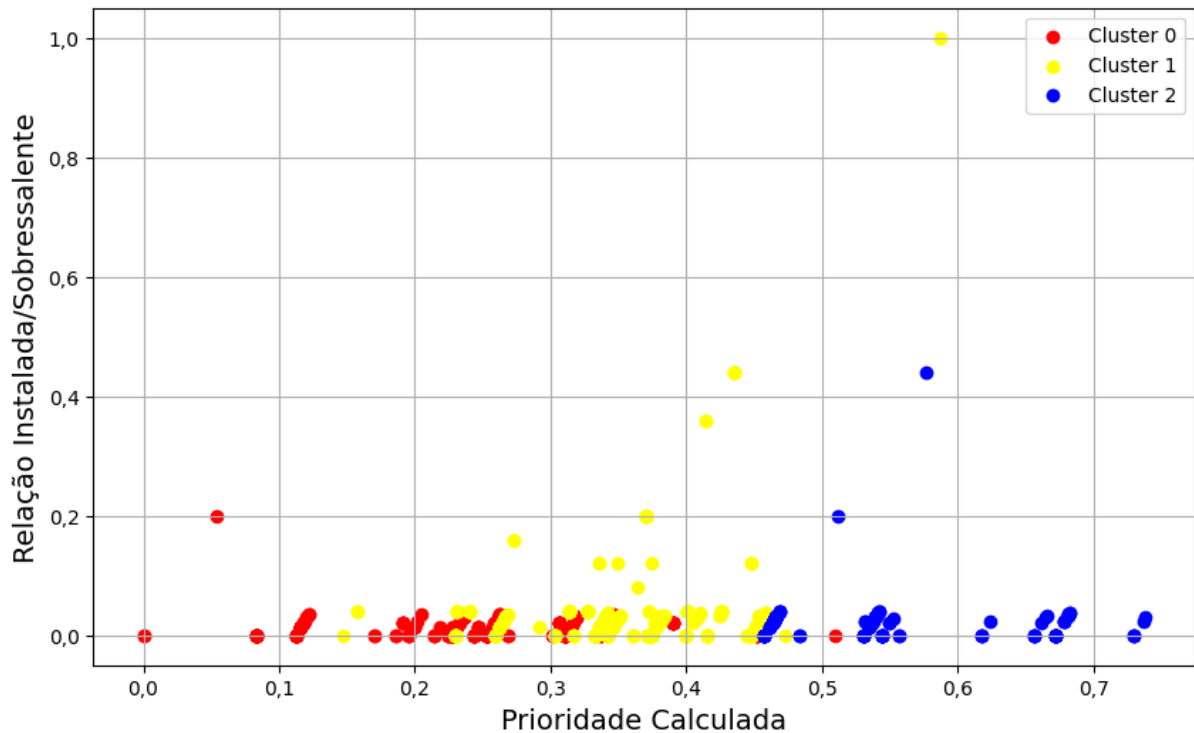
A Figura 15 apresenta a relação entre a quantidade de ativos instalados e a disponibilidade de sobressalentes para cada cluster. Observa-se que o Cluster 2 (azul), que possui a maior prioridade média, exibe uma dispersão significativa na relação instalada/sobressalente. Essa variação sugere que alguns ativos críticos podem não contar com um número adequado de sobressalentes, o que pode representar um risco elevado à operação caso ocorram falhas sem disponibilidade imediata de reposição.

O Cluster 1 (amarelo), com prioridade intermediária, apresenta uma distribuição mais centralizada, mas ainda com certa variabilidade na relação instalada/sobressalente. Isso pode indicar que, embora existam sobressalentes para parte dos ativos desse grupo, alguns itens podem demandar maior atenção na gestão de estoque para evitar impactos na operação.

Já o Cluster 0 (vermelho), composto por ativos de baixa prioridade, possui uma relação instalada/sobressalente mais concentrada em valores reduzidos, sugerindo que há maior disponibilidade de peças de reposição para esses equipamentos. Esse fator pode minimizar os riscos de indisponibilidade desses ativos, que já possuem menor impacto operacional.

Dessa forma, a análise reforça a necessidade de um planejamento estratégico para gestão de sobressalentes, especialmente para os ativos do Cluster 2, que apresentam alta prioridade e maior risco associado à falta de peças de reposição.

Figura 15 – Distribuição da Relacao\_Instalada\_Sobressalente em função da Prioridade Calculada por Cluster.



Fonte: Elaborado pelo autor, 2025.

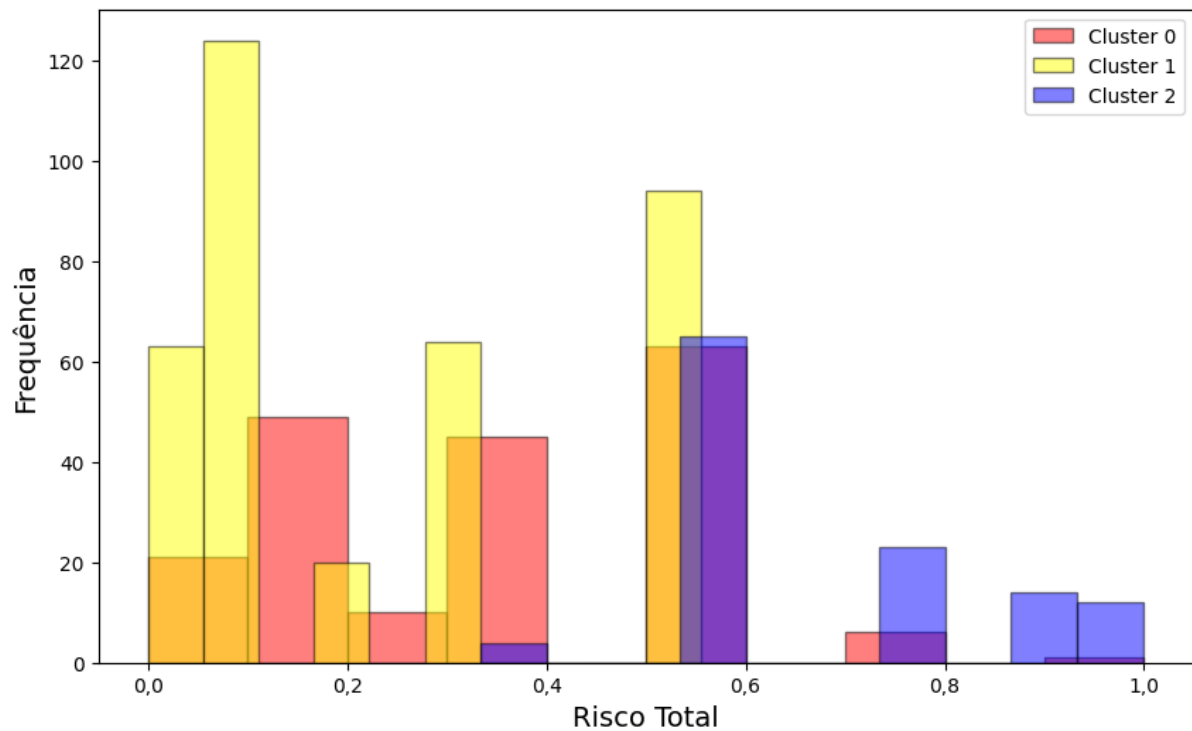
A Figura 16 apresenta a distribuição da variável risco total para cada cluster. Observa-se que o Cluster 2 (azul), composto por ativos de maior prioridade, concentra uma parte significativa dos equipamentos em faixas de alto risco, reforçando a necessidade de estratégias eficazes de manutenção para minimizar impactos operacionais.

O Cluster 1 (amarelo), que representa ativos de prioridade intermediária, apresenta uma distribuição mais ampla, com um grande número de ativos em faixas de baixo a médio risco. Isso indica que, apesar de alguns equipamentos exigirem atenção, a criticidade operacional desse grupo é menos expressiva quando comparada ao Cluster 2.

Já o Cluster 0 (vermelho), formado por ativos de baixa prioridade, apresenta maior frequência em faixas de baixo risco, sugerindo que esses equipamentos demandam menor atenção em termos de manutenção corretiva.

A análise da distribuição do risco total por cluster permite um planejamento mais eficiente das atividades de manutenção preditiva e corretiva, alocando recursos conforme o risco operacional.

Figura 16 – Histograma da distribuição do Risco Total por Cluster.



Fonte: Elaborado pelo autor, 2025.

Embora o valor do ativo não tenha sido utilizado como uma variável de entrada no modelo de clusterização, sua análise permite uma avaliação complementar sobre a classificação dos ativos. A Figura 17 apresenta a média dos valores dos ativos por cluster.

O gráfico indica que o Cluster 0 (baixa prioridade) apresenta o maior valor médio de ativos, seguido pelo Cluster 2 (alta prioridade) e, por último, pelo Cluster 1 (prioridade intermediária). Esse resultado sugere que ativos mais caros não são necessariamente os de maior prioridade, pois a criticidade e o impacto operacional não estão diretamente ligados ao custo do equipamento.

A inversão entre os valores médios do Cluster 0 e do Cluster 2 é um ponto relevante. Isso pode indicar que, embora alguns ativos de baixa prioridade tenham um custo elevado, seu impacto operacional é reduzido, justificando sua menor criticidade dentro da análise de clusterização.

Figura 17 – Valor médio dos Ativos por Cluster.

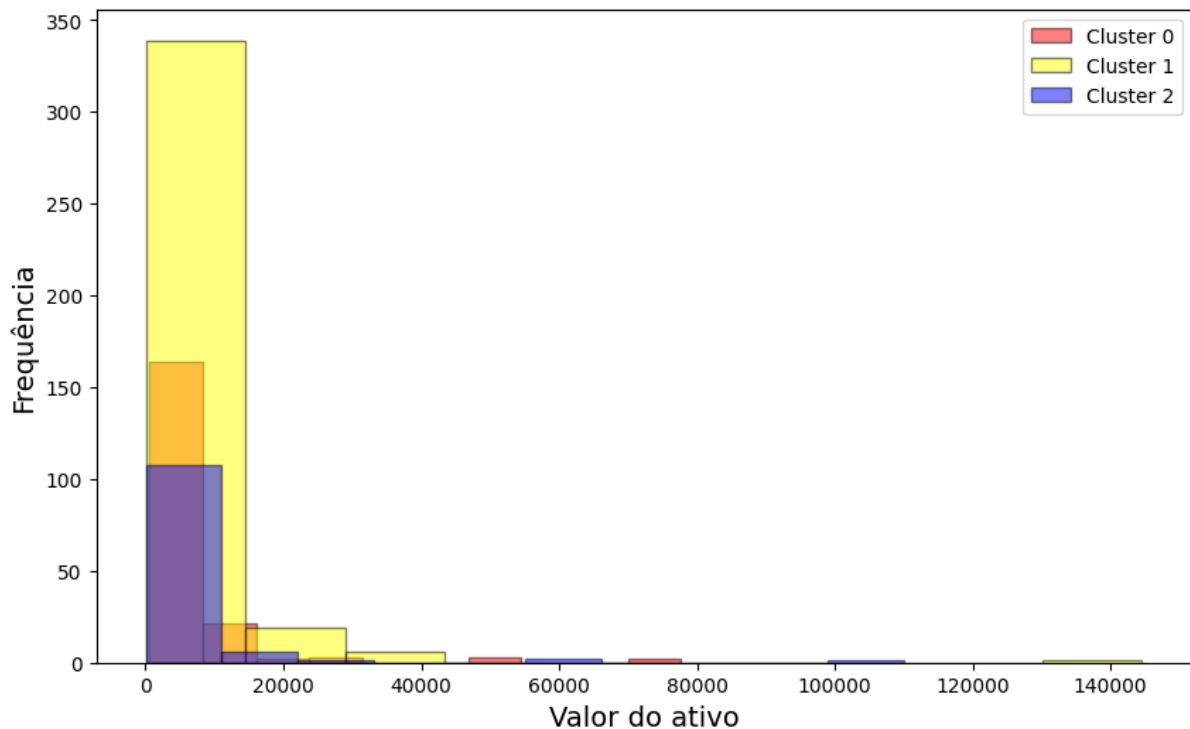


Fonte: Elaborado pelo autor, 2025.

A análise da distribuição dos valores dos ativos em cada cluster (Figura 18) indica que a maioria dos ativos está concentrada em faixas de valores menores, independentemente do cluster.

Entretanto, observa-se que o Cluster 0 apresenta a maior variação no valor dos ativos, indicando que essa categoria contém tanto ativos de alto valor quanto de menor custo. Já o Cluster 1 possui uma distribuição mais concentrada em ativos de menor valor, o que sugere que ativos críticos para a operação não são necessariamente os mais caros, mas sim aqueles cuja falha pode gerar impactos significativos para a planta industrial. O Cluster 2, por sua vez, também inclui ativos de maior valor, reforçando a conclusão de que o custo de um ativo não é o principal fator determinante de sua importância operacional.

Figura 18 – Histograma da distribuição do Valor dos Ativos por Cluster.



Fonte: Elaborado pelo autor, 2025.

A relação entre o valor do ativo e a prioridade corrobora a necessidade de uma gestão que leve em consideração não somente o custo financeiro dos ativos, mas também sua criticidade e risco para a operação, que devem ser integrados conjuntamente para uma gestão orçamentária e financeira eficientes.

## 4.6 Exportação dos Dados e Aplicação Prática

A análise da clusterização neste estudo não se limita apenas à exploração dos dados dentro do ambiente de desenvolvimento, mas também facilita a aplicação prática dos resultados na gestão de ativos de automação. Para isso, foram implementadas duas abordagens principais: a exportação dos dados clusterizados e a criação de uma interface interativa para exploração dos clusters.

Para permitir uma análise posterior mais aprofundada e a integração com outras ferramentas, os dados dos clusters foram exportados para um arquivo Excel, no qual cada cluster identificado pelo K-Means foi salvo em uma aba separada, conforme Figura 19. Isso possibilita que os profissionais responsáveis pela manutenção e gestão possam acessar e manipular as informações de forma intuitiva. O código desenvolvido para essa funcionalidade utiliza a biblioteca pandas e a engine xlsxwriter para a criação do arquivo.

Figura 19 – Arquivo Excel exportado pelo código desenvolvido.

Aplicação	Status do Ciclo de Vida do Ativo Instalado	Valor do ativo	Equipamento	Criticidade	Status do Ciclo de Vida do Ativo Instalado (Numérico)	Criticidade (Numérico)	Risco_Total	Prioridade	Cluster											
2 ISOR DA TVERSORZEE EMANZEE E944	LENZE USHE17241	ATIVO	1	1	USHE17241	71104	71104	10	TS201	1	Não impossível-5-Critico	0	1	0,5	1	1	0,16666667	0,631667	0	
3 P DA MT 2 PLC UR2 NIVERSAL	SIEMENS 100-11A01-	ATIVO	1	1	100-11A01-	5134,12	5134,12	3	MT239	A	1-Raro	Moderac	0	0,666667	0	0,5	0,222222	0,166667	0,267222	0
4 P DA MT 2 PLC 32XDC 217AL D132	SIEMENS 121-1B101-	ATIVO	2	2	121-1B101-	9496,68	9496,68	3	MT239	A	1-Raro	Moderac	0	0,666667	0	0,5	0,222222	0,166667	0,267222	0
5 P DA MT 2 PLC XSCACV032, 2AV	SIEMENS 122-1B100-	ATIVO	2	2	122-1B100-	10280,96	10280,96	3	MT239	A	1-Raro	Moderac	0	0,666667	0	0,5	0,222222	0,166667	0,267222	0
6 P DA MT 2 PLC 30 L1242 + GSD IPI	SIEMENS 242-5D030	ATIVO	1	1	242-5D030	1888,96	1888,96	3	MT239	A	1-Raro	Moderac	0	0,666667	0	0,5	0,222222	0,166667	0,267222	0
7 P DA MT 2 PLC 115F-2-PLI 0/5 MS /	SIEMENS 315-2F14-	ATIVO	1	1	315-2F14-	28313,26	28313,26	4	MT242	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0
8 P DA MT 2VRSORCONTROLN COM IN	SIEMENS 40-11A01-	ATIVO	8	8	40-11A01-	3924,79	31298,32	4	MT242	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0
9 P DA MT 2VRSORONVERTEENAMICS S1	SIEMENS 110-1TE32-	ATIVO	3	3	110-1TE32-	77662,3	232986,78	4	MT242	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0
10 C DA MT 2 PLC VTE 1740NTE 1746-ROCKWELL	1746-P2	FIM DE VIDA	2	2	1746-P2	3151,2	6302,4	3	MT203	B	1-Raro	Moderac	0,33333333	0,33333333	0	0,5	0,222222	0,166667	0,217222	0
11 C DA MT 2 PLC 1747-SN 0 SCANNROCKWELL	1747-SN	FIM DE VIDA	1	1	NARIRAP9	10200	10200	3	MT203	B	1-Raro	Moderac	0,33333333	0,33333333	0	0,5	0,222222	0,166667	0,217222	0
12 C DA MT 2 PLC 1746-OR18NS-SRROCKWELL	1746-OR18	FIM DE VIDA	1	1	1746-OR18	3749,4	3749,4026	3	MT203	B	1-Raro	Moderac	0,33333333	0,33333333	0	0,5	0,222222	0,166667	0,217222	0
13 C DA MT 2 PLC 1746-OR18T ISOLATROCKWELL	1746-OR18	FIM DE VIDA	1	1	1746-OR18	3541,87	3541,8711	3	MT203	B	1-Raro	Moderac	0,33333333	0,33333333	0	0,5	0,222222	0,166667	0,217222	0
14 P DA MT 2 PLC 157322-1D016XDC	SIEMENS 123-1B101-	ATIVO	3	3	123-1B101-	1744,07	5232,21	8	MT243	1	Não impossível-4-Elevad	0	1	0,5	0,75	0,777778	0,166667	0,542778	0	
15 P DA MT 2 PLC 157323-3H8/DO8XDC	SIEMENS 123-1B101-	ATIVO	1	1	123-1B101-	1875,23	1875,23	8	MT243	1	Não impossível-4-Elevad	0	1	0,5	0,75	0,777778	0,166667	0,542778	0	
16 P DA MT 2 PLC 157385-CMENS IMS	SIEMENS 165-0B401-	ATIVO	1	1	165-0B401-	1004,51	1004,51	8	MT243	1	Não impossível-4-Elevad	0	1	0,5	0,75	0,777778	0,166667	0,542778	0	
17 P DA MT 2 PLC 157322-1D016XDC	SIEMENS 123-1B101-	ATIVO	3	3	123-1B101-	1744,07	5232,21	8	MT244	1	Não impossível-4-Elevad	0	1	0,5	0,75	0,777778	0,166667	0,542778	0	
18 P DA MT 2 PLC 157323-3H8/DO8XDC	SIEMENS 123-1B101-	ATIVO	1	1	123-1B101-	1875,23	1875,23	8	MT244	1	Não impossível-4-Elevad	0	1	0,5	0,75	0,777778	0,166667	0,542778	0	
19 P DA MT 2 PLC 157385-CMENS IMS	SIEMENS 165-0B401-	ATIVO	1	1	165-0B401-	1004,51	1004,51	8	MT244	1	Não impossível-4-Elevad	0	1	0,5	0,75	0,777778	0,166667	0,542778	0	
20 ISOR DA TVERSORVER FLEWER FLEX	ROCKWELL:RM002B-F	ATIVO	1	1	RM002B-F	14480	14480	4	MT222	1	1-Raro	1-Elevad	0	1	0,75	0,333333	0,166667	0,365	0	
21 ISOR DA TVERSORVER FLEWER FLEX	ROCKWELL:RM002B-F	ATIVO	1	1	RM002B-F	14480	14480	4	MT222	1	1-Raro	1-Elevad	0	1	0,75	0,333333	0,166667	0,365	0	
22 C DA MT 2 PLC 769-021DC INPUT	ROCKWELL769-0216	ATIVO	6	6	769-0216	1425,05	8550,2983	4	MT222	1	1-Raro	1-Elevad	0	1	0,75	0,333333	0,166667	0,365	0	
23 C DA MT 2 PLC 769-021DC	ROCKWELL769-0216	ATIVO	6	6	769-0216	1425,05	8550,2983	4	MT223	A	1-Raro	1-Elevad	0	0,75	0,333333	0,166667	0,365	0		
24 C DA MT 2 PLC 734-AENAPTER	TWROCKWELL734-AEN1	ATIVO	2	2	1734-AEN1	2933,11	5866,224	4	MT222	1	1-Raro	1-Elevad	0	1	0,75	0,333333	0,166667	0,365	0	
25 C DA MT 2 PLC 1734-IB8/28V DC	ROCKWELL1734-IB8	ATIVO	16	16	1734-IB8	659,95	10559,203	4	MT222	1	1-Raro	1-Elevad	0	1	0,75	0,333333	0,166667	0,365	0	
26 C DA MT 2 PLC 734-0VDC	RELAROCKWELL1734-0V4	ATIVO	12	12	1734-0V4	960,179	11522,149	4	MT222	1	1-Raro	1-Elevad	0	1	0,75	0,333333	0,166667	0,365	0	
27 ISOR DA TVERSORVER FLEWER FLEX	ROCKWELL:RM002B-F	ATIVO	1	1	RM002B-F	14480	14480	4	MT223	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0
28 ISOR DA TVERSORVER FLEWER FLEX	ROCKWELL:RM002B-F	ATIVO	1	1	RM002B-F	14480	14480	4	MT223	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0
29 C DA MT 2 PLC 769-021DC INPUT	ROCKWELL769-0216	ATIVO	6	6	769-0216	1425,05	8550,2983	4	MT223	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0
30 C DA MT 2 PLC 769-021DC	RELAROCKWELL769-0216	ATIVO	6	6	769-0216	1283,93	7703,5697	4	MT223	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0
31 C DA MT 2 PLC 734-AENAPTER	TWROCKWELL734-AEN1	ATIVO	2	2	1734-AEN1	2933,11	5866,224	4	MT223	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0
32 C DA MT 2 PLC 1734-IB8/28V DC	ROCKWELL1734-IB8	ATIVO	16	16	1734-IB8	659,95	10559,203	4	MT223	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0
33 C DA MT 2 PLC 734-0VDC	RELAROCKWELL1734-0V4	ATIVO	12	12	1734-0V4	960,179	11522,149	4	MT223	A	1-Raro	1-Elevad	0	0,666667	0	0,75	0,333333	0,166667	0,311667	0

Fonte: Elaborado pelo autor, 2025.

Além da exportação dos dados, foi desenvolvida uma ferramenta interativa utilizando widgets da biblioteca ipywidgets, que permite a exploração dinâmica dos clusters dentro do ambiente de análise. Essa interface possibilita que os usuários selecionem diferentes clusters e atributos para análise, por meio da escolha das colunas de interesse.

Figura 20 – Exemplo de interação com a interface.

Aplicação	Status do Ciclo de Vida do Ativo Instalado	Valor do ativo	Equipamento	Criticidade	Status do Ciclo de Vida do Ativo Instalado (Numérico)	Criticidade (Numérico)	Risco_Total	Prioridade
1 CLP DA TS 201	DESCONTINUADO	6872,95	TS201	1	0,666667	1,000000	1,000000	0,638333
2 CLP DA TS 201	DESCONTINUADO	254,41	TS201	1	0,666667	1,000000	1,000000	0,638333
3 CLP DA TS 201	DESCONTINUADO	220,00	TS201	1	0,666667	1,000000	1,000000	0,638333
4 CLP DA TS 201	DESCONTINUADO	2475,00	TS201	1	0,666667	1,000000	1,000000	0,638333
5 CLP DA TS 201	DESCONTINUADO	138,57	TS201	1	0,666667	1,000000	1,000000	0,638333
6 CLP DA TS 201	DESCONTINUADO	1313,00	TS201	1	0,666667	1,000000	1,000000	0,638333
7 CLP DA TS 201	DESCONTINUADO	5105,14	TS201	1	0,666667	1,000000	1,000000	0,638333
8 CLP DA TS 201	OBSOLETO	3800,00	TS201	1	1,000000	1,000000	1,000000	0,641667
9 CLP DA TS 201	OBSOLETO	1784,13	TS201	1	1,000000	1,000000	1,000000	0,641667
10 CLP DA TS 201	OBSOLETO	19000,00	TS201	1	1,000000	1,000000	1,000000	0,641667
11 CLP DA TS 201	OBSOLETO	25000,00	TS201	1	1,000000	1,000000	1,000000	0,641667
12 CLP DA TS 201	OBSOLETO	3200,00	TS201	1	1,000000	1,000000	1,000000	0,641667
34 CLP DA MT 242	DESCONTINUADO	11274,01	MT242	A	0,666667	0,666667	0,333333	0,318333
35 CLP DA MT 242	DESCONTINUADO	9593,74	MT242	A	0,666667	0,666667	0,333333	0,318333
36 CLP DA MT 242	DESCONTINUADO	9593,74	MT242	A	0,666667	0,666667	0,333333	0,318333
37 CLP DA MT 242	DESCONTINUADO	431,92	MT242	A	0,666667	0,666667	0,333333	0,318333
38 CLP DA MT 242	DESCONTINUADO	665,63	MT242	A	0,666667	0,666667	0,333333	0,318333
39 CLP DA MT 242	DESCONTINUADO	497,92	MT242	A	0,666667	0,666667	0,333333	0,318333
40 CLP DA MT 242	DESCONTINUADO	2705,42	MT242	A	0,666667	0,666667	0,333333	0,318333
41 CLP DA MT 242	DESCONTINUADO	138,57	MT242	A	0,666667	0,666667	0,333333	0,318333

Fonte: Elaborado pelo autor, 2025.

Com essa implementação, os usuários podem explorar os dados dos clusters de forma interativa, facilitando a gestão dos ativos. Essa ferramenta pode ser útil para priorização de investimentos, planejamento de manutenção e análise de risco dos equipamentos, pois oferece uma visão clara e detalhada sobre os diferentes ativos dentro dos clusters.

## 5 CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho, foi realizada a integração da Inteligência Artificial na gestão de ativos de tecnologia da automação, com o objetivo de auxiliar a tomada de decisão e aprimorar a eficiência operacional. Para isso, foi utilizada a clusterização por meio do algoritmo K-Means, permitindo a segmentação dos ativos com base em variáveis relevantes, como criticidade, risco total, ciclo de vida e relação instalada/sobressalente.

Inicialmente, definiu-se um critério de priorização dos ativos utilizando pesos ajustados por um algoritmo que maximizou o Silhouette Score (0,413), garantindo a coerência na separação dos clusters. O algoritmo identificou os seguintes pesos ideais para a equação de prioridade: 0,26 para o risco total, 0,25 para a criticidade, 0,22 para o status do ciclo de vida do ativo e 0,27 para a relação instalada/sobressalente. Os resultados demonstraram que a abordagem proposta permitiu categorizar os ativos de maneira eficiente, identificando aqueles com maior impacto operacional e necessidade de atenção prioritária.

Na análise dos clusters, verificou-se que os ativos de baixa prioridade apresentaram maior disponibilidade de sobressalentes e menor risco operacional, indicando menor necessidade de ações imediatas. Em contrapartida, os ativos de alta prioridade exibiram menor disponibilidade de reposição e maior criticidade, reforçando a importância de um monitoramento contínuo e estratégias de contingência. Já os ativos de prioridade intermediária apresentaram um equilíbrio entre risco e disponibilidade, sendo candidatos para uma gestão otimizada baseada em manutenção preditiva.

É importante ressaltar que, com a atualização ou entrada de novos dados, a segmentação dos ativos pode sofrer alterações, uma vez que o algoritmo de clusterização recalcula os agrupamentos com base nas novas informações disponíveis. Isso significa que a classificação de um ativo pode mudar conforme variações em seu risco total, disponibilidade de sobressalentes ou outros critérios analisados. Dessa forma, é necessário estabelecer uma gestão contínua, garantindo o monitoramento e a reavaliação periódica dos clusters para embasar a análise de riscos, planejamento financeiro e estratégias de manutenção. Esse acompanhamento dinâmico permite ajustes mais precisos na alocação de recursos e priorização de ações, assegurando que a tomada de decisão seja sempre baseada em dados atualizados e alinhada aos objetivos estratégicos da organização.

Embora o valor do ativo não tenha sido uma variável de entrada na clusterização, sua análise revelou que os ativos de maior valor estão distribuídos entre diferentes clusters, com maior concentração na categoria de média prioridade. Esse resultado indica que nem sempre os ativos mais caros são os mais críticos, reforçando a importância de critérios baseados em impacto operacional e disponibilidade de sobressalentes. No entanto, ativos de alto valor e alta prioridade se destacaram como os mais críticos para a operação, pois além dos elevados custos de substituição, apresentam baixa disponibilidade de reposição, exigindo uma gestão estratégica mais rigorosa.

Além disso, foi desenvolvida uma interface interativa que permite a visualização dos clusters e a filtragem dinâmica dos dados, facilitando a interpretação das informações e auxiliando na tomada de decisão estratégica. A exportação dos resultados em formato Excel viabiliza a integração desses dados em sistemas de gestão de ativos e manutenção preditiva, ampliando o potencial de aplicação da solução desenvolvida.

Com base nos resultados obtidos, conclui-se que a aplicação da Inteligência Artificial na gestão de ativos de tecnologia da automação pode potencializar a gestão financeira e a redução de riscos operacionais, contribuindo para o desenvolvimento das práticas de manutenção e gestão de ativos de automação.

## 5.1 Trabalhos Futuros

Como propostas para trabalhos futuros, sugere-se a exploração e implementação de abordagens que possam aprimorar e expandir os resultados obtidos neste trabalho. As principais direções incluem:

- Integração de um algoritmo de consulta de preços via web: A implementação de um algoritmo que colete automaticamente os valores dos ativos de diferentes fornecedores poderia permitir uma atualização dinâmica dos custos dos equipamentos, melhorando a precisão das análises financeiras e facilitando decisões estratégicas sobre substituição e manutenção;
- Desenvolvimento de modelos baseados no CAPEX para priorização de ativos: A criação de algoritmos que recebam o CAPEX (Capital Expenditure) como entrada e retornem recomendações sobre quais ativos devem ser gerenciados prioritariamente pode otimizar o planejamento financeiro das empresas. No entanto, essa abordagem exige um estudo mais aprofundado sobre a relação entre CAPEX, criticidade dos ativos e impacto operacional, garantindo que a priorização seja fundamentada em critérios financeiros e técnicos bem definidos;
- Aprimoramento do método de clusterização K-Means: O modelo de agrupamento pode ser aprimorado com estratégias como K-Means++ (para melhorar a inicialização dos centroides) e a incorporação de pesos dinâmicos nas variáveis utilizadas, tornando-o mais adaptável a diferentes cenários e variações dos dados ao longo do tempo. Essas melhorias podem aumentar a precisão da segmentação dos ativos e permitir ajustes mais refinados na priorização e gestão.

## REFERÊNCIAS

- ALMEIDA, N. Fundamentos e perspectivas de inovação na gestão de ativos de engenharia. **RAE—Revista de Ativos de Engenharia**, v. 1, n. 1, p. 05–16, 2023. Citado na página 22.
- BALDISSARELLI, L.; FABRO, E. Manutenção preditiva na indústria 4.0. **Scientia cum industria**, v. 7, n. 2, p. 11–22, 2019. Citado 2 vezes nas páginas 16 e 17.
- BARROS, A. **Introdução à técnica de clusterização K-Means**. 2021. <<https://ealexbarros.medium.com/introdução-a-técnica-de-clusterização-k-means-89caf4f56bb3>>. Acesso em: 17 fev. 2025. Citado na página 31.
- BHOLOWALIA, P.; KUMAR, A. Ebk-means: A clustering technique based on elbow method and k-means in wsn. **International Journal of Computer Applications**, Citeseer, v. 105, n. 9, 2014. Citado 3 vezes nas páginas 21, 30 e 31.
- BISHOP, C. M.; NASRABADI, N. M. **Pattern recognition and machine learning**. [S.l.]: Springer, 2006. v. 4. Citado 3 vezes nas páginas 12, 17 e 18.
- BOUSDEKIS, A.; MAGOUTAS, B.; APOSTOLOU, D.; MENTZAS, G. A proactive decision making framework for condition-based maintenance. **Industrial Management & Data Systems**, Emerald Group Publishing Limited, v. 115, n. 7, p. 1225–1250, 2015. Citado 2 vezes nas páginas 23 e 24.
- CARVALHO, T. P. *et al.* A systematic literature review of machine learning methods applied to predictive maintenance. **Computers & Industrial Engineering**, Elsevier, v. 137, p. 106024, 2019. Citado 2 vezes nas páginas 13 e 21.
- COSTA, A. *et al.* Predictive maintenance study for high-pressure industrial compressors: Hybrid clustering models. **arXiv preprint arXiv:2411.13919**, 2024. Citado 3 vezes nas páginas 21, 30 e 31.
- Data At. **Introdução ao Machine Learning - Agrupamento**. 2024. Acesso em: 17 fev. 2025. Disponível em: <<https://dataat.github.io/introducao-ao-machine-learning/agrupamento.html>>. Citado na página 30.
- GROUP, P. **How Artificial Intelligence (AI) Can Be Used in Asset Management**. 2024. Acesso em: 16 fev. 2025. Disponível em: <<https://www.prometheusgroup.com/resources/posts/how-artificial-intelligence-ai-can-be-used-in-asset-management>>. Citado na página 13.
- HENDLER, J. Avoiding another ai winter. **IEEE Intelligent Systems**, IEEE Computer Society, v. 23, n. 02, p. 2–4, 2008. Citado 2 vezes nas páginas 12 e 17.
- HOLLAN, J.; HUTCHINS, E.; KIRSH, D. Distributed cognition: toward a new foundation for human-computer interaction research. **ACM Transactions on Computer-Human Interaction (TOCHI)**, ACM New York, NY, USA, v. 7, n. 2, p. 174–196, 2000. Citado na página 12.
- IKOTUN, A. M. *et al.* K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. **Information Sciences**, Elsevier, v. 622, p. 178–210, 2023. Citado na página 31.
- JAIN, A. K.; DUBES, R. C. **Algorithms for clustering data**. [S.l.]: Prentice-Hall, Inc., 1988. Citado 2 vezes nas páginas 12 e 18.

LIMA, A. L. d. C. D.; ARANHA, V. M.; SPERANDIO, E. G. Manutenção preditiva aplicada a ambientes de missão crítica de supercomputação utilizando inteligência artificial: Uma revisão sistemática de literatura. Centro Universitário SENAI CIMATEC, 2021. Citado na página 17.

LOPES, N.; CAVIQUE, L. Segmentação de clientes e análise dos atributos mais relevantes dos clusters. **Revista de Ciências da Computação**, n. 17, 2022. Citado 2 vezes nas páginas 21 e 29.

MCCARTHY, J.; MINSKY, M. L.; ROCHESTER, N.; SHANNON, C. E. A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. **AI magazine**, v. 27, n. 4, p. 12–12, 2006. Citado 2 vezes nas páginas 12 e 17.

MCKINNEY, W. **Python for data analysis: Data wrangling with Pandas, NumPy, and IPython**. [S.l.]: "O'Reilly Media, Inc.", 2012. Citado na página 27.

MINNAAR, J. R.; BASSON, W.; VLOK, P.-J. Quantitative methods required for implementing pas 55 or the iso 55000 series for asset management. **South African Journal of Industrial Engineering**, South African Institute of Industrial Engineers (SAIIE), v. 24, n. 3, p. 98–111, 2013. Citado na página 13.

OCHI, L. S.; DIAS, C. R.; SOARES, S. S. F. Clusterização em mineração de dados. **Instituto de Computação-Universidade Federal Fluminense-Niterói**, v. 1, p. 46, 2004. Citado na página 19.

OYEWOLE, G. J.; THOPIL, G. A. Data clustering: application and trends. **Artificial Intelligence Review**, Springer, v. 56, n. 7, p. 6439–6475, 2023. Citado na página 21.

PAUL, P. O.; OGUGUA, J. O.; EYO-UDO, N. L. Innovations in fixed asset management: Enhancing efficiency through advanced tracking and maintenance systems. **International Journal of Science and Technology Research Archive**, v. 7, n. 01, p. 019–026, 2024. Citado na página 13.

PINTO, M. C. *et al.* Sistema de manutenção preditiva de falhas em válvulas em um processo industrial utilizando inteligência artificial. Universidade Federal de Uberlândia, 2021. Citado na página 21.

Poli Júnior. **Clusterização de Dados - K-Means**. 2024. Acesso em: 14 mar. 2025. Disponível em: <<https://polijunior.com.br/blog/clusterizacao-de-dados/>>. Citado na página 20.

RAPTIS, T. P.; PASSARELLA, A.; CONTI, M. Data management in industry 4.0: State of the art and open challenges. **IEEE Access**, IEEE, v. 7, p. 97052–97093, 2019. Citado 3 vezes nas páginas 17, 18 e 23.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence: a modern approach**. [S.l.]: Pearson, 2016. Citado 2 vezes nas páginas 12 e 17.

SANTANA, R. V.; PONTES, H. L. J. Aplicação da clusterização por k-means para criação de sistema de recomendação de produtos baseado em perfis de compra. **Navus: Revista de Gestão e Tecnologia**, Serviço Nacional de Aprendizagem Comercial (Senac), n. 10, p. 31, 2020. Citado 2 vezes nas páginas 28 e 35.

SANTOS, B. P.; ALBERTO, A.; LIMA, T. D. F. M.; CHARRUA-SANTOS, F. M. B. Indústria 4.0: desafios e oportunidades. **Revista Produção e Desenvolvimento**, v. 4, n. 1, p. 111–124, 2018. Citado 2 vezes nas páginas 16 e 17.

- SANTOS, F. A. O. *et al.* Utilização de machine learning para prever a recorrência de tumores cerebrais. **Revista Científica da UNIFENAS-ISSN: 2596-3481**, v. 6, n. 5, 2024. Citado na página 28.
- SICHMAN, J. S. Inteligência artificial e sociedade: avanços e riscos. **Estudos Avançados, SciELO Brasil**, v. 35, p. 37–50, 2021. Citado 3 vezes nas páginas 12, 17 e 18.
- STANČIN, I.; JOVIĆ, A. An overview and comparison of free python libraries for data mining and big data analysis. In: IEEE. **2019 42nd International convention on information and communication technology, electronics and microelectronics (MIPRO)**. [S.l.], 2019. p. 977–982. Citado na página 27.
- TESSARINI, G.; SALTORATO, P. Impactos da indústria 4.0 na organização do trabalho: uma revisão sistemática da literatura. **Revista Produção Online**, v. 18, n. 2, p. 743–769, 2018. Citado na página 16.
- TROPIA, C. E. Z.; SILVA, P. P.; DIAS, A. V. C. Indústria 4.0: Uma caracterização do sistema de produção. In: **XVII Congresso Latino-Iberoamericano de Gestión Tecnológica. ALTEC**. [S.l.: s.n.], 2017. Citado na página 16.
- VALENTIM, M. L. P. Estrutura de bases de dados: modelos de metadados e a qualidade de resposta. **Transinformação, SciELO Brasil**, v. 13, p. 67–80, 2001. Citado na página 27.
- VARDAKAS, G.; PAPAKOSTAS, I.; LIKAS, A. Deep clustering using the soft silhouette score: Towards compact and well-separated clusters. **arXiv preprint arXiv:2402.00608**, 2024. Citado na página 21.
- ZUBAIR, M. *et al.* An improved k-means clustering algorithm towards an efficient data-driven modeling. **Annals of Data Science**, Springer, v. 11, n. 5, p. 1525–1544, 2024. Citado na página 20.