



**Universidade do Minho**  
Escola de Engenharia

Solange Fernandes Moreira

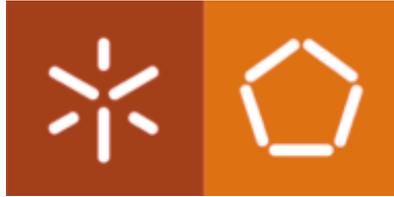
**Deteção de vazamento de gás em sistemas de ar  
condicionado através de técnicas de 6 sigma e  
aprendizagem de máquina**

**Deteção de vazamento de gás em sistemas de ar  
condicionado através de técnicas de 6 sigma e  
aprendizagem de máquina**

Solange Fernandes Moreira

UMinho | 2021

Julho de 2021



**Universidade do Minho**  
Escola de Engenharia

Solange Fernandes Moreira

**Deteção de vazamento de gás em  
sistemas de ar condicionado através de  
técnicas de 6 sigma e aprendizagem de  
máquina**

Julho de 2021



**Universidade do Minho**

Escola de Engenharia

Solange Fernandes Moreira

**Deteção de vazamento de gás em  
sistemas de ar condicionado através de  
técnicas de 6 sigma e aprendizagem de  
máquina**

Dissertação de Mestrado  
Mestrado em Engenharia Industrial

Trabalho realizado sobre a orientação de:  
Professora Dra. Maria Leonilde Rocha Varela e  
Professor Dr. Vaibhav Shah

Julho de 2021

## **Direitos de Autor e Condições de Utilização do Trabalho por Terceiros**

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada. Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

***Licença concedida aos utilizadores deste trabalho***



**Atribuição-Nãocomercial-SemDerivação**

**CC BY-NC-ND**

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

## **AGRADECIMENTOS**

Agradeço a Deus pela oportunidade de desenvolver esse trabalho.

Aos meus pais, que sempre me inspiraram com seus exemplos e direcionamento de vida.

Ao meu esposo Alessandro Trindade e minha filha Louise Trindade, que estão sempre ao meu lado apoiando nas minhas escolhas e decisões.

A empresa que possibilitou o estudo com base nos seus produtos e processos, patrocinando o conhecimento e o desenvolvimento de seus colaboradores.

A Eridanis, empresa francesa, com atuação forte em desenvolvimento de projetos focado na Indústria 4.0 e deu um grande suporte na fase de POC (prova do conceito) desse projeto.

A Universidade do Minho, em Portugal que possibilitaram esse intercâmbio de conhecimentos, em especial meus orientadores, Professora Dra. Maria Leonildes Rocha Varela e Professor Dr. Vaibhav Shah.

A Faculdades IDAAM, no Brasil, com todos os seus professores que compartilharam conhecimento contribuindo para o nosso aprendizado e em especial ao professor M.Sc Vicente Fernandes Tino.

E a todos que me deram apoio para que esse projeto se realizasse.

## **DECLARAÇÃO DE INTEGRIDADE**

Declaro ter atuado com integridade na elaboração do presente trabalho acadêmico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

## **Deteção de vazamento de gás em sistemas de ar condicionado através de técnicas de 6 sigma e aprendizagem de máquina**

### **RESUMO**

Como em toda revolução, onde mudanças bruscas são perceptíveis e a sociedade naturalmente se adapta, aprende e segue em frente, a revolução da Indústria 4.0 vem trazendo as oportunidades com o avanço da tecnologias que podem facilitar, otimizar e melhorar os processos atuais. Esse projeto visa o desenvolvimento de um teste de conceito para um método de deteção de vazamento de gás refrigerante automático, para utilização em processos de manufatura de ar condicionados, a fim de aumentar a confiança em reter produtos defeituosos na fábrica, minimizando a interferência humana.

A metodologia 6 sigma foi utilizada como auxílio no planejamento, execução e análise de experimentos sequenciais, com 4 ciclos de experimentações, envolvendo variáveis de ambiente industrial, produto e uma câmera de infravermelho, responsável por coletar a imagem térmica da área de estudo do vazamento de gás. Métodos de aprendizagem de máquinas supervisionado foram aplicados aos dados de temperatura do produto para treinamento de algoritmos, para classificação de regiões com “Vazamento” e sem vazamento “Normal”. O algoritmo Regressão Logística apresentou melhor performance nas previsões, demonstrando que é possível fazer a deteção de vazamento de gas, de forma automática em ambiente industrial, passando segurança para o projeto seguir para próxima fase de aprimoramento do algoritmo e implementação.

Palavras-chaves: Aprendizagem de máquina; Experimentos Planejados (DOE); Indústria 4.0; Regressão Logística; Testes de vazamento de gás.

# **Gas leak detection in air conditioner cooling system using 6 sigma and machine learning techniques**

## **ABSTRACT**

As in any revolution, where sudden changes happen, and the society naturally adapts, learns and moves on, the Industry 4.0 revolution has brought opportunities with technologies that can facilitate, optimize and improve current processes. This project aims to develop a concept test for an automatic gas leak detection method, to be used in air conditioning manufacturing processes, in order to increase confidence in retaining defective products at the factory, removing human interference.

The 6-sigma methodology was used to aid in the planning, execution and analysis of sequential experiments, with 4 experimental cycles, involving industrial variables, product and a thermal camera, responsible for collecting the thermal image and temperature of the leakage area. Supervised Machine Learning method was used to train algorithms on temperature data set to classify an area as “Gas leakage” and “Normal”. The regression logistic algorithm had the best performance in the predictions, showing that is possible to detect “Gas leakage” area in automatic decision in industrial environmental, giving confidence to forward to next step of the project to improve the algorithm and implementation.

Keywords: Gas leakage test; Industry 4.0; Logistic Regression; Machine Learning; Planned Experiment (DOE)

## ÍNDICE GERAL

<b>Direitos de Autor e Condições de Utilização do Trabalho por Terceiros</b> .....	i
<b>AGRADECIMENTOS</b> .....	ii
<b>DECLARAÇÃO DE INTEGRIDADE</b> .....	iii
<b>RESUMO</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>ÍNDICE GERAL</b> .....	vi
<b>ÍNDICE DE FIGURAS</b> .....	viii
<b>ÍNDICE DE TABELAS</b> .....	xi
<b>LISTA DE ABREVIATURAS, SIGLAS E ACRÔNIMOS</b> .....	xii
1 Introdução.....	1
1.1 Enquadramento.....	1
1.2 Objetivos .....	3
1.3 Metodologia .....	4
1.4 Organização Dissertação .....	4
2 Revisão da literatura.....	6
2.1 Funcionamento do ar condicionado .....	6
2.2 Controle de qualidade na fabricação de um produto .....	8
2.3 Métodos de testes para deteção de vazamento .....	9
2.4 Cameras térmicas - Deteção de vazamento de gás através de imagem térmica.....	13
2.5 Características de vazamento .....	14
2.6 Metodologia para melhoria em produtos e processos – 6 sigma.....	15
2.7 Aprendizagem de máquina “Machine Learning” .....	26
2.8 Software “Orange” .....	32
3 Definição do problema e proposta de solução.....	35
3.1 Definição do problema .....	35
3.2 Proposta de solução.....	37
4 Teste do conceito – PDSA´s.....	40
4.1 PDSA#1 “Plan Do Study Action” .....	42
4.2 PDSA #2.....	49
4.3 PDSA #3.....	54
4.4 PDSA #4.....	56
5 Aprendizagem de máquina para identificação de vazamento .....	61

5.1	Estratégia de amostragem (recolha de dados) .....	61
5.2	Aprendizagem de máquina utilizando software Orange .....	64
6	Resultados e discussões.....	69
7	Conclusões e trabalhos futuros.....	71
8	Referências Bibliográficas .....	73

## ÍNDICE DE FIGURAS

Figura 1 – A energia na forma de calor flui do corpo de maior temperatura para o de menor temperatura (MARQUES & Araujo, 2009). .....	6
Figura 2 – Sistema de refrigeração de ambiente fechado (M. Coutinho, 2016).....	7
Figura 3 - Métodos de testes com pressão positiva e negativa/vácuo (Rottländer et al., 2016). 9	9
Figura 4 – Esquema de teste de vazamento de gas com farejador (Inficon, 2019b).....	10
Figura 5 - Esquema de teste de vazamento de gás com farejador de He (Rottländer et al., 2016). .....	11
Figura 6 – Esquema de teste utilizando camara de He – Adaptado de Inficon, (2019a) .....	11
Figura 7 – Espectro eletromagnético com area de IR (Vollmer, Michael; Mollmann, 2010)..	13
Figura 8 - Ciclo de Deming -Adaptado de Moen et al. (2012). .....	16
Figura 9 - Aprendizado seqüencial: Múltiplos ciclos de Deming – Adaptado de Moen et al. (2012). .....	17
Figura 10 - Diagrama de causa e efeito com termos de experimentação planejada – Adaptada de Moen et al. (2012). .....	18
Figura 11 – Matriz de experimentos com 2, 3 e 4 fatores – Adaptada de Moen et al. (2012). 22	22
Figura 12 – Gráfico de resultados da variável resposta Y.....	25
Figura 13 – Gráficos de efeitos principais T (Tempo) e Te (Temperatura).....	25
Figura 14 – Gráficos de efeitos da interação T (Tempo) x Te (Temperatura).....	26
Figura 15 - Técnicas de aprendizado de máquina e seus dados requeridos (Mohammed et al., 2016).....	27
Figura 16 – Exemplo de visualização de uma árvore de decisão – Adaptada de Côrtes et al. (2002). .....	29
Figura 17 - Dados de treinamento, onde existe um número infinito de hiperplanos que podem separar as classes (Côrtes et al., 2002). .....	30
Figura 18 – Visualização do programa Orange.....	33
Figura 19 – Conjunto de ferramentass de processamento.....	33
Figura 20 – Tela inicial do programa Orange– Fonte: Software Orange.....	34
Figura 21 – Processo de montagem de ar condicionado de janela.....	35
Figura 22 - Pontos de junções de solda em um trocador de calor.....	36
Figura 23 – Gráfico de produtos com vazamento no deposito.....	37

Figura 24 – Imagem de um produto com vazamento de gás refrigerante (Lado esquerdo com câmara de IR e Lado direito com câmara convencional) .....	38
Figura 25 - Fases de desenvolvimento do projeto de teste de vazamento.....	39
Figura 26 – Diagrama de causa e efeito: levantamento de variáveis .....	40
Figura 27 – Tipos e tamanhos de furos que permitem o vazamento de gás.....	41
Figura 28 – Camera Optris PI640 – Fonte: <a href="https://www.optris.com/thermal-imager-optris-pi-640">https://www.optris.com/thermal-imager-optris-pi-640</a> .....	41
Figura 29 - Árvore de amostragem do experimento – DOE#1 .....	43
Figura 30 - Medição da variável resposta Y1 (tempo para reduzir a temperatura em 0,5°C na região do vazamento: 15 segundos) .....	45
Figura 31 - Medição da variável resposta Y2 (tempo para visualizar a imagem do vazamento, seja gás escapando ou mudança de cor permanente: 12,3segundos) .....	45
Figura 32 – Tempo para detecção do vazamento Y1 e Y2 – DOE#1 .....	46
Figura 33 – Gráfico de Efeitos Principais e Interações – DOE#1 .....	47
Figura 34 - Gráficos de efeito principal para Y1 (tempo para reduzir 0,5°C) .....	47
Figura 35 - Gráficos de efeito principal para Y2 (tempo para visualizar a imagem do vazamento) .....	48
Figura 36 - Gráficos de interações de segunda ordem para Y1 (tempo para reduzir 0,5°C) ...	48
Figura 37 - Árvore de amostragem do experimento (DOE#2).....	50
Figura 38 - Gráfico de tempo para detecção do vazamento (DOE#2).....	52
Figura 39- Gráfico de efeitos principais e interações (DOE#2).....	52
Figura 40 - Gráficos de efeito principal para Y1 (tempo para reduzir 0,5°C) – DOE#2 .....	53
Figura 41 - Gráficos de efeito principal para Y2 (tempo para visualizar a imagem do vazamento) .....	53
Figura 42 - Gráficos de interação AB para Y1 (tempo para reduzir 0,5°C) – DOE#2 .....	53
Figura 43 - Árvore de amostragem do experimento (DOE#3).....	56
Figura 44 - Árvore de amostragem do experimento (DOE#4).....	58
Figura 45 – Gráfico de tempo para detecção do vazamento (DOE#4) .....	59
Figura 46 – Gráfico de efeitos principais e interações (DOE#4) .....	60
Figura 47 – Gráficos de efeito principal para Y1 (tempo para reduzir 0,5°C) – DOE#4.....	60
Figura 48 – Gráficos de efeito principal para Y2 (tempo para visualizar a imagem do vazamento) – DOE#4.....	60
Figura 49 – Árvore de amostragem – DOE#4.....	62

Figura 50 – Pontos de coleta de temperatura através da imagem de IR.....	62
Figura 51 – Estratégia de amostragem de TREINO e TESTE / Amostras de temperatura para cada experimento.....	63
Figura 52 – Projeto de aprendizado de máquina para classificação de vazamento.....	65
Figura 53 – Carregamento de dados de um arquivo Excel e ajustes do tipos de dados e funções .....	66
Figura 54 – Visualização da tabela com dados carregados .....	66
Figura 55 – Visualização gráfica dos dados utilizando a ferramenta “Line Plot” .....	67
Figura 56 – Algoritmos de aprendizagem conectados ao teste de performance “Testand Score” .....	68
Figura 57– Resumo dos PDSAs para Y (tempo para detectar vazamento).....	69

## ÍNDICE DE TABELAS

Tabela 1– Tipos de variáveis, adaptado de (Reis, 2002).....	19
Tabela 2 - Quantidade de testes para experimentos fatoriais $2k$ .....	22
Tabela 3– Níveis dos fatores T e Te.....	23
Tabela 4 - Matriz de experimentação com resultados (Y) e cálculo dos efeitos.....	23
Tabela 5– Matriz de confusão de classificação binária (Kohavi, R; Provost, 1998) .....	31
Tabela 6 - Fatores e níveis – DOE#1 .....	42
Tabela 7 - Variáveis de ruído – DOE#1 .....	43
Tabela 8 - Matriz de experimentação e coleta de dados – DOE#1 .....	44
Tabela 9 - Fatores e níveis para o DOE#2.....	50
Tabela 10 - Variáveis de fundo ou ruído e métodos de controle.....	50
Tabela 11 - Matriz de experimentação e coleta de dados (DOE#2).....	51
Tabela 12 - Fatores e níveis para o DOE#3.....	55
Tabela 13 - Variáveis de fundo ou ruído e métodos de controle.....	55
Tabela 14 - Fatores e níveis para o DOE#4.....	57
Tabela 15 - Variáveis de fundo ou ruído e métodos de controle.....	57
Tabela 16 - Matriz de experimentação e coleta de dados (DOE#4).....	58
Tabela 17 - Dados de Treino (amostra dos dados).....	64
18 – Tabela com total de dados para treino e teste dos algoritmos .....	64
Tabela 19 - Avaliação dos algoritmos com dados de Treino e Teste.....	70
Tabela 20 - Matriz de confusão dos algoritmos .....	71

## **LISTA DE ABREVIATURAS, SIGLAS E ACRÔNIMOS**

CA	Classification Accuracy (Acurácia na Classificação)
C&E	Causa e Efeito
DOE	Design of Experiments (Experimentos Planejados)
FN	Falso Negativo
FP	Falso Positivo
I4.0	Indústria 4.0
IR	Infrared Radiation (Radiação Infravermelha)
ML	Machine Learning (Aprendizagem de Máquina)
PDSA	Plan/Do/Study/Act
SVM	Support Vector Machine
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

# 1 Introdução

Neste capítulo será apresentado o enquadramento sobre o tema estudado da dissertação assim como os objetivos, a metodologia utilizada para a realização deste trabalho e a organização do relatório de dissertação.

## 1.1 Enquadramento

Esse estudo foi conduzido numa empresa de ar condicionado, inserida no Polo Industrial de Manaus – PIM, situado no coração da Amazônia Ocidental, um dos mais modernos centros industriais e tecnológicos em toda a América Latina, reunindo atualmente mais de 600 indústrias de ponta nos segmentos Eletroeletrônico, Duas Rodas, Naval, Mecânico, Metalúrgico e Termoplástico, entre outros, que geram mais de meio milhão de empregos diretos e indiretos (SUFRAMA, n.d.)

O ar condicionado, um aparelho de uso doméstico e industrial, cujo objetivo é controlar a temperatura e umidade de um ambiente fechado é o resultado de várias pesquisas e tentativas de solucionar problemas afetados pela condição ambiente, ao longo dos anos. Foi em 1902 que Willis Carrier, um engenheiro de 25 anos formado pela Universidade de Cornell, nos EUA, inventou um processo mecânico para condicionar o ar. O controle do clima finalmente foi colocado em prática, para resolver o problema de uma empresa de impressão em dias quentes de Nova York, onde foi feita a primeira instalação por Carrier (Web Arcondicionado, 2019). Desde então, o ar condicionado vem sofrendo constantes evoluções tecnológicas, viabilizando seu uso doméstico, industrial, comercial e automotivo, trazendo conforto térmico ao ambiente, de forma sustentável, com uso de gases menos nocivos, sendo o maior desafio dos fabricantes, a redução de consumo de energia.

No final de 2016, de 1,6 bilhões de unidades de ar condicionado residencial instaladas no mundo, a China possui a maior quantidade, os Estados Unidos em segunda posição e o Japão em terceira posição, sendo que os 2 primeiros são responsáveis por um pouco mais de 50% do total. O Brasil ocupava a 9ª posição. É o que mostra o relatório da International Energy Agency (International Energy Agency, 2018). Com esse aumento de demanda no mundo o desafio para garantir a qualidade do produto também cresce nas fábricas de ar condicionados. A qualidade do produto está entre os KPIs da empresa, que segundo (Paramenter, 2007), KPIs representam um conjunto de medidas com foco nos aspectos de performance organizacionais que são mais críticos para o sucesso atual e futuro da organização. O KPI da qualidade representa a taxa de

falha oriundas de intervenções para conserto de produtos em garantia, onde o seu acompanhamento direciona projetos de melhorias para reduzir as ocorrências de intervenções, melhorando a qualidade do produto e a relação com o consumidor ao experimentar o produto.

Quando ocorre a falha “Não Refrigera” ou “Não Aquece” no ar condicionado, uma das razões está relacionada a falta de gás refrigerante, um dos componentes responsáveis pela troca de calor com o ambiente que se deseja refrigera/aquecer, que escapou por algum furo ou micro furo presentes no sistema de refrigeração. Para evitar que uma falha indesejável de um produto chegue até o cliente, testes e inspeções fazem parte do controle de qualidade no processo das empresas, para garantir as principais características de entrega do produto. No processo de fabricação do ar condicionado, testes para detectar a presença de vazamento no produto são aplicados, porém uma das dificuldades para se identificar a presença da falha “Vazamento de gás” está relacionada a invisibilidade do gás refrigerante. As técnicas e métodos adotados atualmente para detecção de vazamento de gases refrigerantes, nas empresas de Ar Condicionado estão relacionados a testes de pressão, detectores de presença de gases, tanques de água, cabine hermeticamente fechadas.

Como os sistemas de detecção de vazamento atual possui uma dependência significativa de equipamentos e processamento humano, e portanto vulnerabilidade de se deixar passar produtos com vazamento para o consumidor final, perspectiva-se que essa dependência pode ser minimizada através do uso de processamento de imagens ou dados aplicando técnicas de aprendizagem de máquina para detecção automática de vazamento.

No contexto da Indústria 4.0 (ou I4.0), uma das principais características centra-se na digitalização de ativos físicos e integração de sistemas digitais, onde interfaces digitais possuem um papel preponderante na comunicação. Para (Ribeiro, 2017) um dos habilitadores digitais permite converter elementos físicos em informações digitais para subsequente tratamento. Nesse sentido, espera-se com esse projeto conseguir o benefício deste tipo de avanço providenciado pela I4.0, de modo a transformar a imagem capturada de uma câmara termográfica, traduzindo em presença ou não de vazamento de gás, através de aprendizagem em padrões de imagens ou dados e posteriormente tomada de decisão.

Visto que o ar condicionado tem em seu sistema a presença de um gás refrigerante circulando, uma alternativa a ser estudada é a termografia, que segundo (Filho et al., 2006) é uma técnica de análise térmica de um corpo ou sistema, sem contato físico. Para (Jadin & Ghazali, 2014) é uma das ferramentas que pode ser utilizada para inspeção e sua tecnologia de *Infrared* permite a medição de temperatura do objeto sem contato com ele, além de atender uma

larga escala de temperatura, resposta rápida de leitura e alta precisão. Dessa forma, o fenômeno “Vazamento de gás” pode ser explorado através de estudos de padrões de dados coletados pela câmera com a aplicação de algoritmos de aprendizagem de máquina.

Por se tratar de uma nova tecnologia ainda não aplicada em processos de Ar Condicionado, um teste inicial do conceito foi recomendado para análise de viabilidade do projeto.

Para validar o conceito, foram utilizadas ferramentas de 6 sigma, uma metodologia desenvolvida pela Motorola em 1987, que se baseia num conjunto de ferramentas estatísticas e da qualidade utilizadas para melhoria contínua de produtos e processos. Uma das principais ferramentas utilizadas para testar uma proposta ou solução são Experimentos Planeados, que segundo (Moen et al., 2012) trata de uma mudança feita a um produto ou processo, de maneira controlada com o propósito de aprendizagem. Para os experimentos desse estudo serão considerados variáveis de ambiente industrial, produto e da câmara térmica utilizada para medição do vazamento.

Após a viabilidade de detecção do vazamento de gás pela câmara, as imagens de produtos com vazamentos e sem vazamentos serão processadas e utilizadas de forma inteligente para tomada de decisão.

## **1.2 Objetivos**

### **Objetivos Gerais**

O objetivo geral deste trabalho consiste na realização de um estudo de um teste automático para detecção de vazamento de gás refrigerante de sistemas de ar condicionado no processo produtivo, através da metodologia de 6 sigma e aprendizagem de máquina, a fim de minimizar a subjetividade humana do teste atual.

### **Objetivos específicos:**

1 - Avaliar a possibilidade de detectar vazamento de gás em produtos de ar condicionado, utilizando câmera de termográfica IR em um ambiente industrial;

2 - Utilizar os dados capturados pela câmera térmica para aprendizagem de máquina e avaliar sua habilidade em classificação de regiões com “Vazamento” e sem vazamento “Normal”.

### 1.3 Metodologia

O estudo é de natureza exploratória, com estratégia de Investigação e ação, que segundo (C. P. Coutinho et al., 2009): “Pode ser descrita como uma família de metodologias de investigação que incluem ação (ou mudança) e investigação (ou compreensão), utilizando um processo cíclico”. Este ciclo alterna entre a ação e reflexão/compreensão de forma contínua até que o conhecimento seja gerado.

Para validação do conceito, técnicas de experimentos fatoriais planejados (DOE) foram utilizadas, auxiliando na definição da estratégia de amostragem, levando em consideração a manipulação dos fatores, seus repetitivos níveis e efeitos no resultado dos testes:

- Processo: *Takt time* da linha, condições ambientais, tipo de gás, pressão de gases
- Câmara termográfica: Configurações de funcionamento
- Produto: Design de tubos, tipo de tubo, tamanho dos furos

A câmera térmica possibilita a captura da imagem ou vídeo com IR, bem como a temperatura em cada pixel da imagem/vídeo ao longo do tempo. Os dados de temperatura nas regiões de vazamento e fora da região de vazamento foram extraídos através de um software que acompanha a câmera (USB IR CAM Rel. 3.9.3064.0) e realizado o tratamento para sua utilização no treino de algoritmos de aprendizagem de máquina, a fim de desenvolvimento de habilidades em identificar padrões de regiões com “Vazamento” de gás e regiões sem vazamento de gás “Normal”. Esse estudo foi feito com o auxílio do software “Orange”, um programa que trabalha com aprendizagem de máquina (*machinelearning*) e visualização de dados (*data visualization*). Trata-se de um software de código aberto, que pode ser utilizado por usuários iniciantes e especialistas, pois o software possui uma grande variedade de ferramentas de análise de dados interativa de visualização, exploração, pré-processamento e modelagem de dados, de forma amigável (LARHUD, 2018).

### 1.4 Organização Dissertação

O presente trabalho de dissertação e relatório foi elaborado com base nas principais fases subjacentes:

**Fase 1** - Revisão bibliográfica: Teorias sobre o funcionamento de ar condicionado, metodologia 6 sigma, métodos de testes para detecção de vazamento de gases, tipos de vazamento, câmeras termográficas IR, aprendizagem de máquina e ferramentas para aplicação.

**Fase 2** - Situação atual do problema e proposta conceitual de solução: Levantamento da condição atual de um processo de manufatura de ar condicionado, focando na falha vazamento de gás refrigerante, considerando as variáveis de produto, processo e da câmera termográfica.

**Fase 3** - Teste do conceito: Planejamento, execução e análise de experimentos sequenciais para avaliar se é possível detectar vazamento de gás utilizando uma câmera termográfica em um processo produtivo de ar condicionados.

**Fase 4** – Aprendizagem de máquina: Avaliação do uso dos dados da câmera térmica para aprendizado de algoritmos a fim de identificar padrões de comportamento na temperatura da superfície do trocador de calor do ar condicionado, para identificar se uma região tem ou não vazamento de gás. Para esse estudo, foi utilizado o software “Orange”.

**Fase 5** – Resultados, discussões, conclusões e recomendações para trabalhos futuros.

## 2 Revisão da literatura

Neste capítulo é apresentada uma breve revisão de literatura sobre o funcionamento de um ar condicionado, como é feito o controle de qualidade em seu processo produtivo e os métodos de detecção de vazamento conhecidos, bem como a literatura sobre o uso e aplicação de câmeras termográficas. Adicionalmente, a teoria sobre 6 sigma e seus métodos de experimentação foram também usados, bem como a literatura sobre aprendizagem de máquinas e sua utilização através do software “Orange”.

### 2.1 Funcionamento do ar condicionado

Para entender o funcionamento do ar condicionado, recorreu-se a alguns princípios da física, tais como a troca de calor entre 2 corpos. Segundo (MARQUES & Araujo, 2009), James Joule, entre 1840 e 1849, realizou medições bastante precisas sobre a equivalência mecânica do calor por diversos métodos e confirmou, experimentalmente, que calor é uma forma de energia. Ressalta também que é possível variar a energia interna de um sistema fornecendo ou retirando energia. Quando a transferência de energia ocorre exclusivamente devido a uma diferença de temperatura entre o sistema e a vizinhança, a energia transferida recebe o nome de calor, ilustrado pela figura 1.



Figura 1 – A energia na forma de calor flui do corpo de maior temperatura para o de menor temperatura (MARQUES & Araujo, 2009).

Esse é o princípio básico usado pelo ar condicionado, onde o ar presente no ambiente a ser refrigerado é direcionado para entrar em contato com uma peça interna ao aparelho de ar condicionado (Evaporador), cuja temperatura é inferior a temperatura ambiente. Pelo princípio da troca de calor, remove o calor do ar ambiente e o mesmo retorna ao ambiente com uma temperatura mais baixa. Esse ciclo se repete até que ocorra o equilíbrio térmico entre o ar ambiente e a temperatura do Evaporador, o qual pode ser melhor visualizado na figura 2.

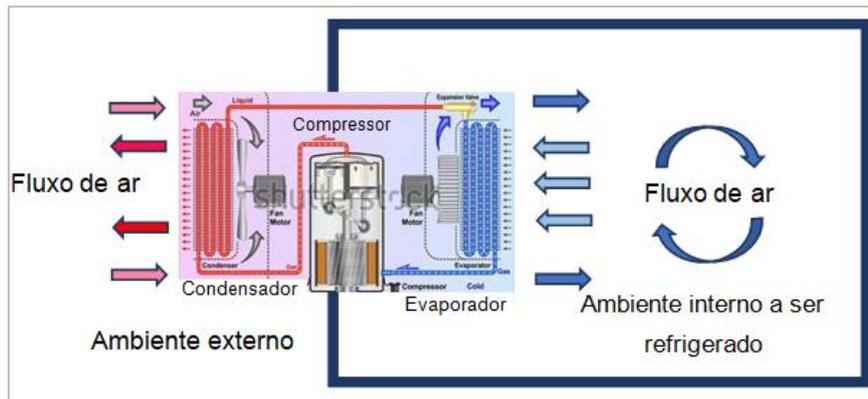


Figura 2 – Sistema de refrigeração de ambiente fechado (M. Coutinho, 2016).

Os principais elementos do ar condicionado responsáveis pelo circuito de refrigeração, proporcionando troca de calor, encontram-se na figura 2 e suas respectivas funções estão na sequência, com base nas explicações de um sistema de refrigeração feitas por Creder, (2004):

- *Gás refrigerante*: Fluido que percorre um circuito hermeticamente fechado, variando de estado gasoso para líquido e vice-versa, com temperaturas baixas e elevadas, provocadas pelo circuito de refrigeração, proporcionando o fenomeno da troca de calor entre dois ambientes;
- *Válvula de expansão de gás (Capilar)*: O gás em estado líquido que sai do condensador, passa por uma restrição, chamada de válvula de expansão de gás, transformando o gás de líquido para gasoso e temperatura baixa;
- *Evaporador*: O gás, sob a forma de vapor úmido, chega frio no Evaporador, que tem a função de promover a troca de calor com o ambiente quando o gás passa por ele a uma temperatura baixa, absorvendo o calor do ambiente interno;
- *Compressor*: O refrigerante, sob forma de vapor saturado, chega ao compressor, que tem a função de comprimir o gás refrigerante e provocar mudança de temperatura no gás, que entra no compressor frio e sai quente, em forma de vapor super aquecido, com alta pressão;
- *Condensador*: Tem a função de promover a troca de calor com o ambiente externo quando o gás refrigerante, vindo do compressor, pelo condensador a uma temperatura alta, condensando o gás que se torna líquido. O ambiente externo deve estar a uma temperatura mais baixa que do trocador de calor, absorvendo o calor do Condensador;
- *Sistema forçado de ventilação*: Esse sistema trabalha em paralelo ao sistema de refrigeração, promovendo a ventilação forçada do ar, passando pelos trocadores de calor (Condensador e Evaporador) e devolvendo ao ambiente interno e externo.

## 2.2 Controle de qualidade na fabricação de um produto

No controle de qualidade é necessário considerar o conceito subjacente ao termo qualidade e seus aspectos, que segundo (Ishikawa, 1990, p. 16), quando é referente a qualidade de um produto, entende-se que seja “a qualidade que a pessoa irá comprar com satisfação”. Portanto, fica a responsabilidade das empresas de desenvolver, planejar, produzir e entregar um produto ou serviço que atenda as expectativas do consumidor.

Existem quatro aspectos da qualidade que são consideradas para satisfazer o consumidor, segundo Ishikawa, (1990):

Q – Qualidade: Performance, durabilidade, dimensões, aparência, embalagens, etc;

C – Custo: Características que afetam o custo e preço: Custo unitário, produtividade, matéria-prima, lucro, etc;

D – Entrega: Características relacionadas a quantidade e tempo de entrega: Volume de produção, volume de vendas, inventário, etc;

S – Serviço: Área preparada para possíveis problemas que aparecem depois do produto ser vendido, tais como período de garantia, segurança, peças de reposição, facilidade de reparo, etc.

De acordo com Ishikawa, (1990), a qualidade de um produto pode ser garantida por vários métodos:

- 1) Através de Inspeção - Inspeção 100%, Inspeção amostral, verificações ou inspeção autônoma;
- 2) Através do processo - Controle do processo, estudos de capacidade de processo;
- 3) Durante o desenvolvimento de novos produtos.

A maioria das empresas adotam os métodos acima, onde a Inspeção é aplicada nas características de qualidade definidas e denominadas como “Itens de controle”, que segundo (Werkema, 1995) são características mensuráveis por meio das quais um processo é gerenciado para que os resultados indesejáveis sejam tratados e seus processos melhorados.

O processo de ar condicionado, do ponto de vista de qualidade, possui elementos comuns a qualquer outro processo, com etapas de construção/montagem do produto e inspeções em Itens de controle que possam afetar diretamente a satisfação do consumidor.

## 2.3 Métodos de testes para detecção de vazamento

Rottländer et al.(2016) consideram dois métodos de testes para detectar vazamento de gás, relacionando pressão interna do objeto, divididos em pressão negativa ou pressão positiva, e tamanho de vazamento mínimo a ser detectado, conforme figura 3.

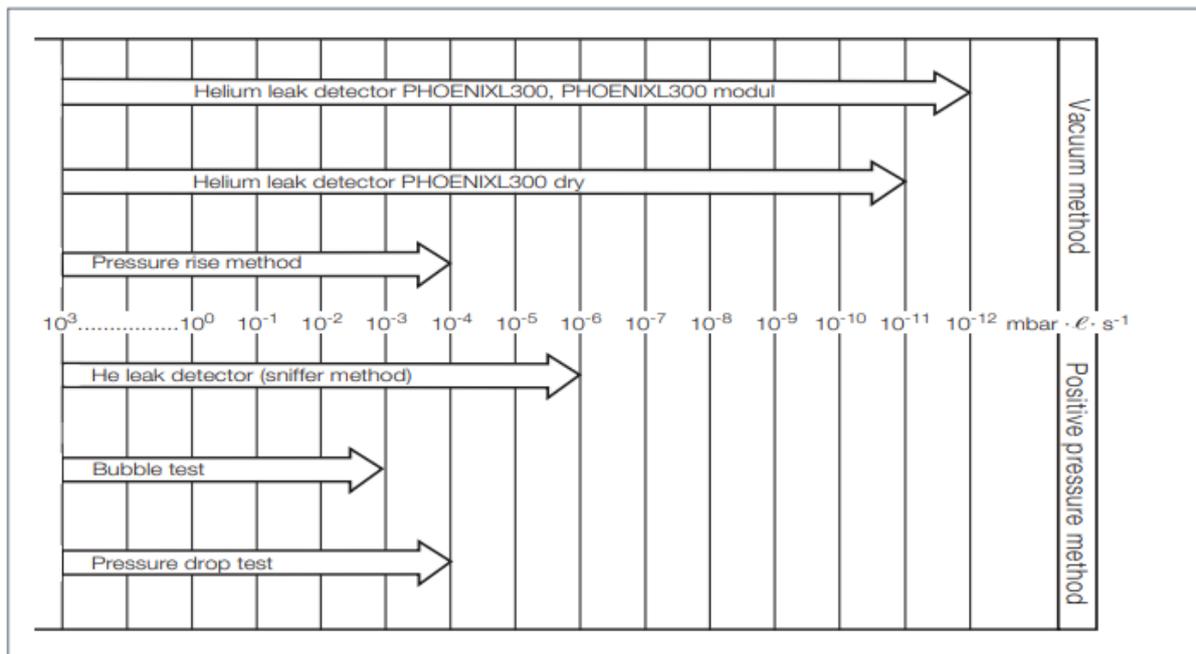


Figura 3 - Métodos de testes com pressão positiva e negativa/vácuo (Rottländer et al., 2016).

Será dado um foco maior na categoria de pressão positiva por se tratar dos métodos mais utilizados na indústria de ar condicionados. Esse método consiste na inserção de um gás dentro do objeto a ser avaliado, com uma pressão maior que a pressão atmosférica (1013mbar). Após a inserção gás, é possível observar a presença ou não do vazamento através de queda de pressão ou através de equipamentos especiais, conhecidos como farejadores, que detectam se o gas que foi injetado escapou para fora do objeto.

Os gases mais comuns utilizados são o ar comprimido para grandes vazamentos, gases halogênios/refrigerantes e o gás hélio (He) para vazamentos muito pequenos.

Para Rottländer et al.(2016), a sensibilidade do método e a precisão da localização de vazamentos por farejadores dependem de:

- Tipo de unidade de farejador usada (ponta do farejador + linha),
- Tempo de resposta do detector de vazamento usado,
- Velocidade de rastreamento e
- Distância da ponta do farejador à superfície do objeto de teste.

### Detetor de gás refrigerante– Inficon HLD 5000

O método que utiliza um detector de gás farejador, requer a aplicação de um gás refrigerante dentro do objeto a ser avaliado (objeto em teste), a uma pressão positiva, onde uma ponta do equipamento detetora do gás (farejador) é aproximado das regiões passíveis de apresentar vazamento do gás, conforme figura 4. Caso seja identificado a presença do gás por fora do objeto, o equipamento sinaliza através de um visor ou sinal sonoro.



Figura 4 – Esquema de teste de vazamento de gas com farejador (Inficon, 2019b).

O funcionamento interno do modelo HLD 5000 é apresentado por Rottländer et al.(2016), onde o equipamento usa a propriedade física das moléculas, capazes de absorver a radiação infravermelha. O gás de teste, capturado pela linha do farejador, flui através de um cubeta que é exposta à radiação infravermelha. Os gases ativos infravermelhos (refrigerante) dentro do testador absorvem uma parte da radiação infravermelha e modifica o sinal infravermelho primário. O sinal infravermelho modificado é detectado por um sensor, processado e exibido. O limite de detecção é em torno de  $5 \cdot 10^{-5} \text{ mbar} \cdot \text{l/s}$ .

### Detetor de gás He – Inficon P 3000

Esse método é similar ao anterior, onde a diferença principal está no tipo de gás de teste. Rottländer et al., (2016) relatam que o método de detecção de gás farejador, o qual identifica vazamento com espectrômetro de massa é o mais utilizado na indústria, devido a sua alta sensibilidade. É possível detectar todos os gases usando espectrometro de massa, no entanto, o hélio é um gás que se mostrou mais apropriado ao uso em testes industriais, devido a sua praticidade, inerte, não explosivo, não toxico, econômico, dentre outras razões.

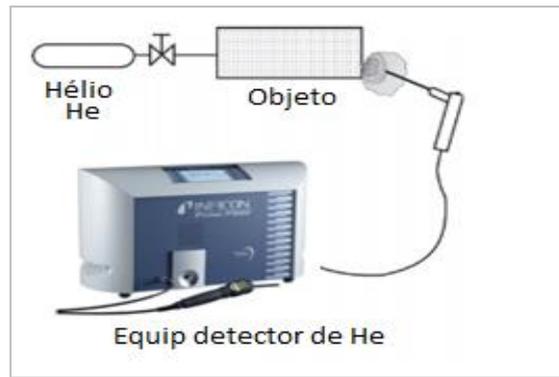


Figura 5 - Esquema de teste de vazamento de gás com farejador de He (Rottländer et al., 2016).

A figura 5 ilustra os elementos necessários para realizar o teste, onde o gás de teste hélio é injetado no interior do objeto, a uma pressão positiva. Um equipamento que detecta a presença de hélio percorre com um farejador por caminhos e junções passíveis de apresentar vazamento, em busca do hélio que possa ter escapado do objeto para fora. Geralmente esse processo é realizado em ambiente aberto e conduzido por uma pessoa/operador que deve seguir o percurso de detecção, respeitando as condições que garantem a efetividade do teste. Para o modelo P 3000 da Inficon o menor vazamento fica em torno de  $1 \cdot 10^{-7}$  mbar  $\cdot$  l / s. A limitação relativa à sensibilidade para a detecção de hélio é devida principalmente à quantidade natural de hélio presente na atmosfera.

### Camara de hélio - Espectometria de massa

A Camara de hélio pode ser feita com pressão positiva ou negativa dentro do produto/objeto submetido ao teste. (Rottländer et al., 2016) ressalta que o material da camara pode ser rígido ou flexível.

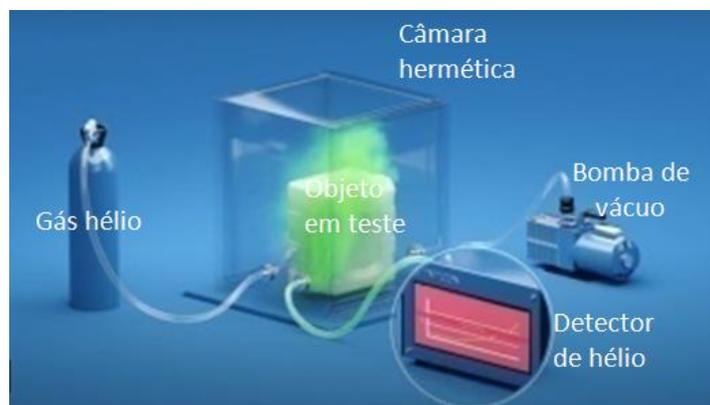


Figura 6 – Esquema de teste utilizando camara de He – Adaptado de Inficon, (2019a)

Considerando pressão positiva, ilustrada na figura 6, o objeto de teste precisa estar com gás hélio em seu interior e submetido a uma câmara selada, que está conectada a um sistema de vácuo e também ao medidor do gás, no caso detector de hélio. Durante o teste, o gás hélio flui pelo vazamento do produto/objeto submetido ao teste e acumula-se dentro da câmara, o qual é detectado pelo medidor da presença do gás conectado a câmara. Caso seja detectado vazamento, o gás hélio precisa ser removido da câmara para garantir o próximo teste. Dependendo do tamanho do produto/objeto e quantidade a ser testada, esse método requer uma grande área, tornando-se uma desvantagem.

Vinogradov et al.(2016) apresentam o teste aplicado com pressão negativa, onde o produto/objeto precisa estar sob vácuo e o gás hélio na parte externa ao produto e dentro da câmara. Nesse caso, o hélio vai para dentro do produto e o equipamento detector de hélio deve estar conectado ao produto.

Como os custos sempre são levados em consideração em processos industriais, para manter o consumo do hélio o mais baixo possível, Rottländer et al.(2016) mencionam que é possível utilizar uma mistura de hélio e ar como gás de teste em vez de 100% de hélio.

### **Estanqueidade por queda de pressão**

Esse é um dos testes de média complexidade na sua implementação, indicado para detecção de vazamentos médios e grandes. A parte interna do produto é conectada a um detector de vazamento de sistema pneumático S9. No interior, o produto é alimentado com ar comprimido. Em seguida, o produto é isolado do sistema que fornece ar. Espera-se que a pressão do ar inserido se mantenha. Aguarda-se alguns segundos e caso aconteça queda de pressão, o resultado do teste indica presença de vazamentos (Vinogradov et al., 2016). Sua principal desvantagem é não se saber qual a posição do vazamento.

### **Estanqueidade por pressão imerso em água – Teste de Bolhas**

Vinogradov et al.(2016) descrevem o método de teste de bolhas como um dos testes mais fáceis de implementar e de baixo custo, pois precisa apenas de um gás inserido no objeto/produto em teste a uma pressão positiva. Esse objeto é submerso em um tanque contendo um líquido. O gás flui através do vazamento do objeto/produto produzindo bolhas no líquido e observada por um operador.

O gás de teste pode ser ar comprimido e o líquido pode ser água. O fluxo mínimo de vazamento detectável para o ar neste método é limitado a  $10^{-3} \dots 10^{-4} \text{ mbar} \cdot \text{l/s}$

## 2.4 Cameras térmicas - Detecção de vazamento de gás através de imagem térmica

Para entender melhor sobre a aplicação de cameras térmicas, se faz necessário alguns conceitos básicos de termografia e radiação infravermelha. Para Marques & Araujo (2009), termografia é um procedimento que permite mapear as diferentes temperaturas ou energia térmica emitida pela superfície de um corpo ou de uma região. Começou a ser usada para fins militares e investigação espacial. Na década de 50, esse procedimento já era usado na medicina e medicina veterinária como ferramenta de auxílio a diagnóstico. Exemplo de uso na medicina veterinária, aplicada em diagnóstico de problemas circulatórios, permitindo a localização de infecções que alteram a distribuição térmica tópica devido a irrigação sanguínea do tecido. Também já é bastante utilizada na engenharia civil, mecânica, automotiva, aeronáutica e eletrotécnica.

Segundo apresentado por Vollmer e Mollmann (2010) cada corpo com uma temperatura acima de zero absoluto ( $-273,15 \text{ }^\circ\text{C} = 0 \text{ Kelvin}$ ) emite radiação eletromagnética de sua superfície, que é proporcional à sua temperatura intrínseca. Uma parte dessa chamada radiação intrínseca é a radiação infravermelha, que pode ser usada para medir a temperatura do corpo. Portanto, a radiação infravermelha intrínseca aos corpos é um princípio básico que possibilita a observação/medição da temperatura da superfície desses corpos ou regiões. Esses são os princípios básicos utilizados pelas cameras térmicas que transformam a radiação infravermelha do corpo em sinal proporcional a sua temperatura.

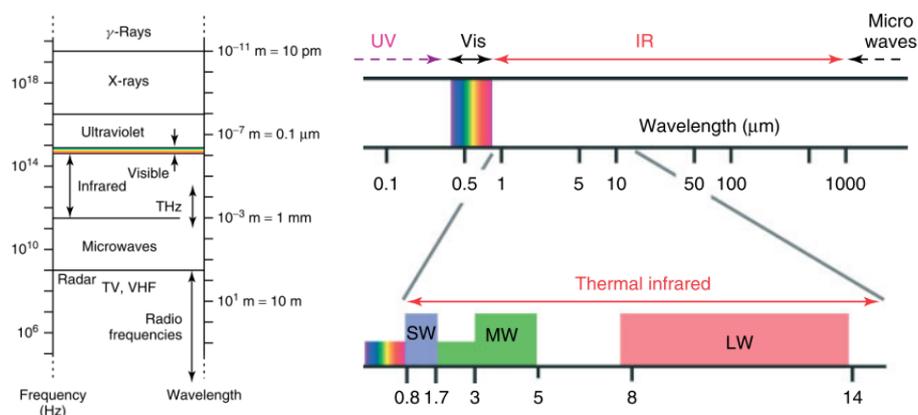


Figura 7 – Espectro eletromagnético com área de IR (Vollmer, Michael; Mollmann, 2010)

A radiação infravermelha faz parte do espectro de radiação eletromagnética, que varia com o comprimento de onda, apresentado na figura 7. Foi descoberta acidentalmente por William Herschel em 1800, ao realizar experimentos sobre o poder de aquecimento e iluminação dos raios coloridos, do espectro de cores visíveis, onde percebeu que a cor vermelha era a mais quente e que havia temperaturas mais elevadas além da cor vermelha, uma região de raios não visível aos olhos humanos, conforme descrito em (Oliveira & Silva, 2014).

Para melhor entendimento, Vollmer e Mollmann (2010), destacam que as cores visíveis possuem comprimento de onda que variam de 400nm a 780nm e a área de infravermelho vem na sequência de 780nm a 1mm. Por esse motivo deu-se o nome de infravermelho. A região para a medição de temperatura infravermelha possui comprimentos de onda que variam de 0,7 a 14  $\mu\text{m}$ . Acima desses comprimentos de onda o nível de energia é tão baixo, que os detectores não são sensíveis o suficiente para identificá-los.

Abaixo temos algumas vantagens da medição de temperatura sem contato:

- Medições de temperatura de movimento ou super-aquecimento de objetos em ambientes normais ou em ambientes perigosos;
- Resposta de leitura muito rápida;
- Medição não interativa, sem influência sobre o objeto de medição
- Medição não destrutiva
- Durabilidade no ponto de medição, sem desgaste mecânico

Não foi encontrado literatura sobre uso de câmeras térmicas para identificação de vazamento de gás em processos industriais de ar condicionado.

## **2.5 Características de vazamento**

O vazamento, segundo informação em Rottländer et al. (2016), é definido como sendo uma "abertura" em uma parede ou barreira através da qual sólidos, líquidos ou gases podem entrar ou sair indesejadamente. Destaca também alguns tipos, tais como vazamentos em conexões destacáveis (flanges, tampas); Vazamento em conexões permanentes (cordões de solda e solda, juntas coladas); Vazamentos devido à porosidade de componentes fundidos, dentre outros.

Nenhum dispositivo ou sistema a vácuo ou sob pressão pode ser absolutamente estanque ao e ele realmente não precisa ser, ressaltam Rottländer et al. (2016). O importante é que a taxa de vazamento seja abaixo da pressão de operação necessária.

Para registrar vazamentos quantitativamente, o conceito de "taxa de vazamento" com o símbolo  $q_L$  e o a unidade mbar l / s foi introduzida. Uma taxa de vazamento de  $q_L = 1 \text{ mbar} \cdot \text{l} / \text{s}$  está presente quando em um recipiente fechado e evacuado com um volume de 1l (litro) da pressão sobe 1 mbar por segundo ou, quando houver pressão positiva no recipiente, a pressão cai em 1 mbar.

## **2.6 Metodologia para melhoria em produtos e processos – 6 sigma**

A melhoria contínua em produtos e processos é uma diretriz básica de empresas que buscam melhores resultados no ambiente competitivo e para se manter no mercado. Nessa busca incessante, o 6 sigma, um conceito desenvolvido na década de 80 pela Motorola e posteriormente aplicado na General Electric (GE) com grandes resultados, proporcionou uma maneira consistente de comparar o desempenho da empresa com os requisitos do cliente, com uma meta ambiciosa de qualidade perfeita (PANDE et al., 2000). É um método direcionado a redução de variações em processos, com o intuito de entregar um produto ao consumidor com uma taxa de falha mínima, na ordem de 3,4 defeitos por um milhão de produtos produzidos, reduzindo custos, melhorando produtividade a qualidade.

Uma definição de seis sigma, segundo PANDE et al. (2000), trata de um sistema abrangente e flexível para alcançar, sustentar e maximizar o sucesso do negócio, direcionado pelo entendimento próximo das necessidades do cliente, uso de fatos, dados e análise estatística para gerenciar, melhorar e reinventar os processos de negócios. Suas origens e ferramentas são encontradas nos ensinamentos de grandes pensadores da “qualidade”, tais como W. Edwards Deming e Shewhart. Um dos princípios teve como base a contribuição de dois conceitos desenvolvidos por Shewhart: Controle Estatístico de Processo (CEP), que considera a variação natural de um processo repetitivo de  $\pm 3$  sigma e o conceito de que o controle deveria direcionar melhorias considerando um fluxo de Especificação => Produção => Inspeção. Esse fluxo sofreu evoluções por Deming e deu origem ao ciclo PDSA Plan-Do-Study-Act, conhecido como Ciclo de Deming e também denominado por ele como Ciclo de Shewhart para aprendizado e melhoria (Moen & Norman, 2009). Portanto, o 6 sigma direciona uma cultura de redução de variação com o uso de ferramentas estatísticas associadas a uma cultura de aprendizado e melhoria contínua.

A última versão do ciclo de PDSA é apresentado por Moen et al. (2012), onde combinam alguns componentes que aceleram a aquisição de conhecimento para a aplicação de ações que

resultam em melhoria, de forma estruturada: Modelo de melhoria (3 perguntas fundamentais), Ciclo PDSA e Método Científico.

Segundo Langley et al. (2009), três perguntas fundamentais compreendem o modelo de melhoria, que direcionam o trabalho prévio para propor mudanças em processos, produtos ou sistemas, com aplicação de forma ampla para qualquer tipo de estudo, estimulando o aprendizado:

1 O que estamos tentando realizar?

Objetivo do esforço; Benefícios; Foco do projeto

2 Como saberemos se uma mudança é uma melhoria?

Base do aprendizado; Dados para mostrar a diferença entre o antes e depois; Medição que representa a melhoria

3 Quais mudanças podem ser feitas que resultarão em uma melhoria?

O que deve ser alterado no processo/produto/sistema; Sugestões podem vir do conhecimento atual, tecnologias ou soluções ainda desconhecidas.

Moen et al. (2012) consideraram alguns elementos do método científico dentro do ciclo PDSA, figura 8, tais como perguntas, previsões e planejamento de testes (P), comparar resultados obtidos com previsões (S), reflexão sobre o aprendizado com os testes e próximos passos (A).



Figura 8 - Ciclo de Deming -Adaptado de Moen et al. (2012).

Com a execução de vários ciclos de PDSA, torna-se claro o processo iterativo do conhecimento em construção, saindo de uma condição de baixo conhecimento sobre o assunto até a sua convergência para resposta da pesquisa em andamento, ilustrado pela figura 9, o aprendizado sequencial com a aplicação de múltiplos Ciclos.

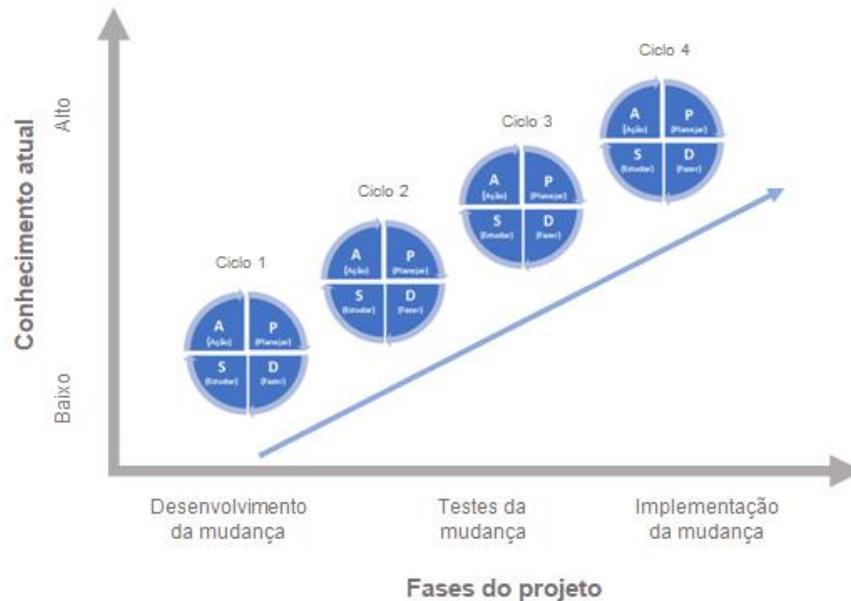


Figura 9 - Aprendizado sequencial: Múltiplos ciclos de Deming – Adaptado de Moen et al. (2012).

A execução de múltiplos ciclos possibilita o aprendizado sequencial, aumentando gradativamente o grau de confiança nas mudanças que realmente farão a diferença no resultado, que atendam as necessidades das organizações e dos consumidores.

Moen et al. (2012) sugerem três princípios básicos de testes que devem ser levados em consideração:

- Teste em baixa escala e construção do conhecimento de forma sequencial;
- Coletar dados ao longo do tempo;
- Incluir amplas condições nos testes sequenciais;

### Experimentos planejados

As técnicas de experimentação utilizadas para gerar o conhecimento possuem uma grande importância na aplicação do ciclo PDSA, visto que o seu resultado direciona os próximos passos, sejam para rever a teoria, testar novamente, abandonar o estudo ou implementar as ações que se mostraram significativas para concluir o estudo.

Para Moen et al. (2012) experimentos planejados consistem numa mudança feita em um processo ou produto, de maneira controlada com o propósito de aprendizado. Para facilitar o entendimento de como funciona essa técnica de experimentação e a linguagem adotada, utiliza-se do mesmo conceito do diagrama de Causa & Efeito, conforme figura 10.

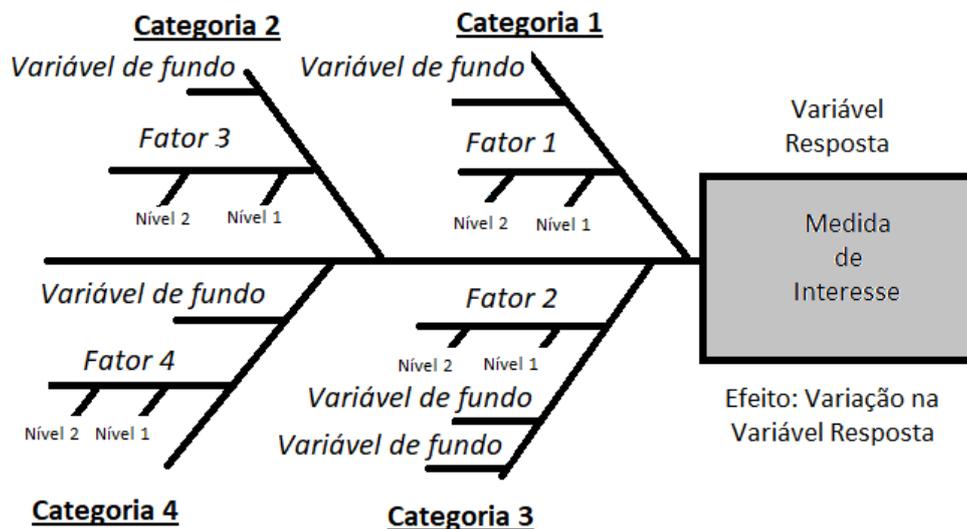


Figura 10 - Diagrama de causa e efeito com termos de experimentação planejada – Adaptada de Moen et al. (2012).

Do lado esquerdo encontram-se as Causas ou Fatores, onde as mudanças no processo/produtos serão consideradas para estudo. Do lado direito, temos o Efeito, representado pela Variável Resposta, que pode variar em função dos fatores/causas modificados de forma controlada.

Quando se conhece qual/quais fatores, que através de testes, variando seus níveis provocam uma variação significativa na variável resposta, temos o conhecimento se concretizando.

É possível ainda ser representado por uma equação linear:  $Y = f(x) + \text{Ruído}$ . Os  $x$ 's são as variáveis de entrada, presentes em um processo, produto ou sistema. Encontram-se do lado esquerdo do diagrama de C&E, que ao sofrerem mudanças, podem provocar uma variação no resultado  $Y$ . O  $Y$  ou  $Y$ 's são as variáveis de saída de um processo, produto ou sistema, geralmente estão associados a alguma característica de qualidade mais próxima do consumidor. Encontram-se do lado direito do diagrama de C&E e sofrem influencia direta da variação das variáveis de entrada  $x$ 's. Ruídos são as variáveis presentes no experimento que não foram levadas em consideração por desconhecimento.

## Termos e definições em experimentações

Os termos utilizados aqui são baseados no trabalho de Fisher, que utilizou pela primeira vez em estudos de agricultura, em busca de melhorar a performance das plantações, em 1933.

**Variável Resposta (Y):** É a medida de resultado observada nas amostras do experimento e também conhecida como variável dependente. Geralmente representa o objetivo do estudo, associada a características de qualidade de um produto, processo ou sistema. Pode ser classificada como quantitativa ou qualitativa. Importante que apresente valores em uma escala numérica que faça sentido, para possibilitar o cálculo da magnitude dos efeitos.

Tabela 1– Tipos de variáveis, adaptado de (Reis, 2002)

	Tipo de variáveis			
	Quantitativas		Qualitativas ou categóricas	
	Discretas	Contínuas	Nominais	Ordinais
Definição	Características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores. Valores inteiros. Geralmente são o resultado de contagens.	Características mensuráveis que assumem valores em uma escala contínua, para as quais valores fracionados fazem sentido. Geralmente devem ser medidas através de algum instrumento.	Não existe ordenação dentre as categorias.	Existe uma ordenação entre as categorias.
Exemplos	Número de filhos, número de bactérias por litro de leite, número de cigarros fumados por dia, número de doentes por semana, número de falhas por mês	Peso (balança), altura (régua), tempo (relógio), pressão arterial, idade, temperatura (termômetro).	Sexo, cor dos olhos, fumante/não fumante, doente/sadio.	Estágio da doença (inicial, intermediário, terminal), mês de observação (janeiro, fevereiro, dezembro), nível de satisfação (baixo, médio, alto)

Para as variáveis qualitativas apresentadas na tabela 1, verifica-se que não há uma escala numérica que possa ser medida como variável resposta. Nesses casos, uma alternativa para o estudo é utilizar uma escala tipo Likert, transformando essa variável qualitativa em uma escala numérica ordinal que possa representar e dar sentido a medição. Por exemplo, para a variável “Estágio da doença” que pode assumir resultados “Inicial, Intermediário e Final”, pode-se adotar uma escala de 1 para Inicial, 2 para Intermediário e 3 para Final.

**Fator (x):** Conhecida como variável independente ou causal, pode ser ajustada de maneira controlada em um experimento e observado o seu impacto na variável resposta. Os fatores podem ser definidos na construção do diagrama de causa e efeito, alocando-os do lado esquerdo (causas), juntamente com as teorias dos especialistas no assunto que servem de base para serem

considerados no estudo. Podem ser do tipo qualitativo (fornecedores, máquinas, material) e quantitativo (pressão, quantidade, comprimento, temperatura).

**Variável de fundo (x):** Chamada de variável contextual, que pode afetar o resultado do experimento, porém não varia da mesma forma que um fator no experimento. Se identificada na fase de planejamento, pode ser levada em consideração a forma que será tratada, tais como manter constante, bloco ou tomando nota. Exemplos de variáveis de fundo são: tempo, operador, instrumento, cavidades dentro de um molde.

**Variável de ruído:** É uma variável onde seu potencial efeito não foi considerado na variável resposta durante o planejamento do experimento. Somente observada durante a execução ou análise do experimento. Podem estar associadas a alguma causa especial, como por exemplo o setup do experimento, temperatura ambiente, umidade. Formas de minimizar seu impacto seria a randomização.

**Níveis:** São os ajustes onde cada fator ou variável de fundo será testada. Para um fator qualitativo, Fornecedor A e Fornecedor B são os níveis para o fator Fornecedores. Para um fator quantitativo, o fator Temperatura será ajustado nos 70°C e 90°C.

**Efeitos:** É a magnitude da mudança ou impacto provocado na variável resposta quando o fator ou variável de fundo mudam de um nível para outro. O efeito pode ser linear com impacto de fatores independentes ou interações de fatores;

## **Experimentos fatoriais**

Experimentos fatoriais é uma das estratégias de experimentação, composta por um plano de testes estruturado, levando em consideração os fatores a serem estudados e seus respectivos níveis, as variáveis de fundo e a variável resposta de interesse, a fim de promover o aprendizado numa relação de causa e efeito. Nesse tipo de experimento todos os fatores são estudados ao mesmo tempo, com possibilidades de quantificar o seus respectivos efeitos na variável resposta, bem como o efeito de interações entre fatores na variável resposta. A estratégia de experimentação é também conhecida como “*Design of Experiments*” DOE (Montgomery & Wiley, 2013).

Existem dois tipos de experimentos fatoriais: *Fatorial Completo* e *Fatorial Fracionado* que auxiliam na definição da quantidade de amostras do experimento. O Fatorial completo é mais adequado quando se tem poucos fatores (até 4) e o Fatorial Fracionado é apropriado para experimentos com uma quantidade grande de fatores (a partir de 5), onde a quantidade de amostras é definida com uma fração do Fatorial Completo. Abaixo será dado mais detalhes sobre Fatorial Completo.

**Fatorial completo:**, consiste em todas as combinações possíveis de fatores e seus níveis. Por exemplo, se três fatores são estudados, Fator A com 2 níveis, Fator B com 3 níveis e Fator C com 5 níveis, a quantidade de combinações possíveis de serem testadas são 30 ( $2 \times 3 \times 5$ ) (Moen et al., 2012).

Como o fatorial completo é sensível a quantidade de fatores e níveis, para facilitar a aplicação na Indústria, levando em consideração que vários ciclos de PDSA serão conduzidos, utilizar fatores em 2 níveis é suficiente para promover o conhecimento. Portanto, o planejamento de experimento aqui será considerado 2 níveis para cada fator, com a notação fatorial abaixo:  $2^k$  onde **k** representa a quantidade de fatores a serem estudados e **2** a quantidade de níveis de cada fator.

Considerando um estudo com 3 fatores, teremos um plano de teste de  $2^3$  que resulta em 8 combinações possíveis entre fatores e níveis.

Para facilitar a visualização e comunicação das combinações de fatores e níveis Moen et al., (2012) mostram a construção de matrizes do experimento, figura 11, utilizando uma convenção comum para os dois níveis, adotando ( - ) para o nível inferior e ( + ) para o nível superior. Esta convenção facilitará o cálculo dos efeitos e pode ser adotada para variáveis contínuas ou qualitativas. Abaixo temos exemplo de experimentos com 2, 3 e 4 fatores e 2 níveis ( - e + ) para cada fator.

Teste	Fator 1	Fator 2
1	-	-
2	+	-
3	-	+
4	+	+

$2^2 = 4$  combinações  
4 testes

Teste	Fator 1	Fator 2	Fator 3
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+

$2^3 = 8$  combinações  
8 testes

Teste	Fator 1	Fator 2	Fator 3	Fator 4
1	-	-	-	-
2	+	-	-	-
3	-	+	-	-
4	+	+	-	-
5	-	-	+	-
6	+	-	+	-
7	-	+	+	-
8	+	+	+	-
9	-	-	-	+
10	+	-	-	+
11	-	+	-	+
12	+	+	-	+
13	-	-	+	+
14	+	-	+	+
15	-	+	+	+
16	+	+	+	+

$2^4 = 16$  combinações  
16 testes

Figura 11 – Matriz de experimentos com 2, 3 e 4 fatores – Adaptada de Moen et al. (2012).

É possível notar na figura 11 que a cada inclusão de 1 fator no estudo, a quantidade de testes dobra de tamanho, seguindo um padrão em função do fatorial completo  $2^k$  que pode ser visto também na tabela 2, onde experimentos com quantidade de fatores a partir de 5 pode se tornar proibitivo em função da quantidade elevada de testes necessários. Para minimizar e reduzir os recursos necessários, existe outro tipo de experimento *fatorial fracionado*, onde a matriz é uma fração do fatorial completo.

Tabela 2 - Quantidade de testes para experimentos fatoriais  $2^k$

$2^k$	Quantidade de fatores (k)	1	2	3	4	5	6	7	8	9	10
	Total de testes ( $2^k$ )		2	4	8	16	32	64	128	256	512

Todos os elementos apresentados até agora estão relacionados ao planejamento (P do PDSA): Fatores, Níveis, Variáveis de fundo, variáveis resposta, quantidade de testes, matriz de experimentação, recursos, etc.

Na seqüência, encontra-se a fase de realização dos testes (D do PDSA), com o auxílio da matriz de experimentos, fazendo os ajustes de cada nível dos fatores, observando o resultado do respectivo teste.

**Exemplo:** Uma fábrica de bolos tem recebido reclamações sobre a qualidade dos bolos vendidos, alguns estavam ligeiramente queimados e outros estavam assado parcialmente. O chefe da cozinha suspeita de dois fatores que podem impactar no resultado da qualidade do bolo: Tempo de assar (T) e a Temperatura (Te) do forno. Um experimento foi planejado conforme abaixo.

*Fatores:* 2 fatores - (T) Tempo e (Te) Temperatura

*Variável de resposta:* Como a reclamação dos clientes é uma variável qualitativa, esta pode ser representada por uma escala numérica de 1 a 10 (tipo Likert), onde tendendo para 1 a qualidade do bolo é ruim e tendendo para 10 a qualidade do bolo é ótima. Um degustador vai provar amostras de bolo e dar uma nota para o sabor classificando-os de 1 a 10, conforme escala abaixo:

(1, 2) Muito ruim (3,4) Ruim (5,6) Neutro (7, 8) Bom (9,10) Muito bom

*Níveis dos fatores:* Os fatores estão com dois níveis. Duas condições de ajuste

Tabela 3– Níveis dos fatores T e Te

Fator	Nível	
	( - )	( + )
(T) Tempo	20min	30min
(Te) Temperatura	160	180

*Matriz de experimentação:* Essa matriz é construída a partir do fatorial  $2^k$ , onde temos 2 fatores e 2 níveis para cada fator, resultando em um total de 4 combinações possíveis (testes), gerando a matriz de experimentação abaixo:

Tabela 4 - Matriz de experimentação com resultados (Y) e cálculo dos efeitos

		Níveis					
		( - )	20	160			
		( + )	30	180			
Teste	Ordem de testes	T (min)	Te (°)	T x Te	Y	(sabor 1 a 10)	
1	1	-	-	+	2		
2	2	+	-	-	8		
3	3	-	+	-	9		
4	4	+	+	+	3		
Efeito		0	1	-6			

*Coleta de dados:* A tabela 4 mostra 4 testes, onde para cada teste foi feito um bolo com sua respectiva receita, considerando os mesmos ingredientes e alterando os fatores Tempo (T) e Temperatura (Te) nos níveis planejados. Cada bolo foi submetido a uma avaliação de sabor/degustação e atribuído uma nota. Exemplo, para o teste 1 o Tempo (T) para assar o bolo foi de 20min, a Temperatura (Te) foi de 160°C e o resultado do sabor foi uma nota 2.

Observa-se uma coluna de interação T x Te, onde esta resulta da multiplicação das colunas T e Te. Vale ressaltar que os ajustes de níveis durante o experimento, só são realizados nas colunas T e Te. O aprendizado sobre a interação é adquirido sem recursos adicionais.

*Análise dos dados:* Após a realização dos 4 bolos com suas respectivas receitas, a coluna Y é preenchida com a avaliação dos bolos (notas de 0 a 10) e a análise dos dados seguem em busca de entender se os fatores T e Te produzem um efeito significativo na variável resposta (Y: sabor do bolo).

Antes de ir direto aos cálculos, é importante um olhar prático nos dados, com perguntas mais concretas: O experimento gerou variação? Uma variação de 2 a 9 no exemplo do bolo faz sentido? Tivemos bolos bons e ruins?

Os dados podem ser representado de forma gráfica e quantitativa, considerando os dados crus, efeitos principais e suas interações (T, Te e TxTe):

Para encontrar o efeito de um fator principal sobre variável resposta, basta calcular a média dos resultados quando o fator está ajustado no nível (+) e subtrair da média dos resultados quando o fator está ajustado no nível (-). Os resultados dos efeitos encontram-se na tabela 4.

$$\text{Ex: Efeito (T)} = (8+3)/2 - (2+9)/2 = 5,5 - 5,5 = 0$$

$$\text{Efeito (Te)} = (9+3)/2 - (2+8)/2 = 6 - 5 = 1$$

*Análise prática:* A figura 12 mostra os resultados de cada teste, o qual se observa que obteve-se bolos com sabor ruim (Teste 1 e 4) e bolos com sabor bom (8 e 9). Um resultado prático bom, pois o experimento gerou variação significativa. Mas quais os fatores são responsáveis por um resultado de bolo ruim ou bolo bom?

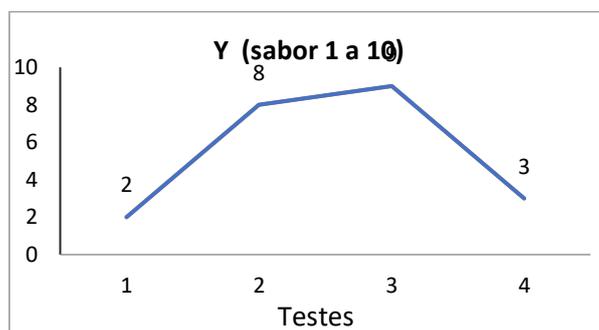


Figura 12 – Gráfico de resultados da variável resposta Y

*Análise gráfica e quantitativa.* A figura 13 mostra que a magnitude do efeito de T quando muda do nível (-) para o nível (+) é zero e a magnitude do efeito de Te quando muda do nível (-) para o nível (+) é 1. Se os efeitos foram tão baixos, por que tivemos bolos com uma variação tão grande?

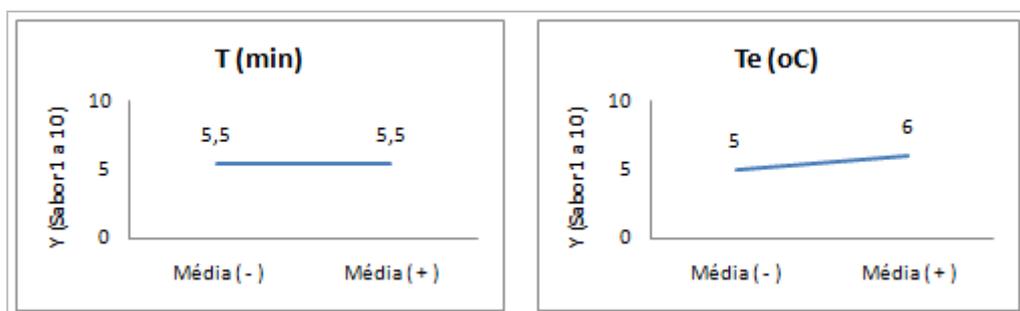


Figura 13 – Gráficos de efeitos principais T (Tempo) e Te (Temperatura)

A figura 14 (lado esquerdo) mostra que a magnitude do efeito da interação TxTe quando muda do nível (-) para o nível (+) é de -6, nesse caso o sinal negativo significa o sentido da inclinação da reta. Quando comparada com os efeitos principais, a interação TxTe provoca uma variação significativa na variável resposta sabor do bolo. Quando a interação está ativa no experimento, significa que a variação na variável resposta depende dos níveis em que os fatores se encontram.

O gráfico do lado direito mostra a interação dos fatores TxTe, onde os valores de Y próximo de 10 podem ser atingidos em duas situações:

Situação 1: Quando o tempo (T) assumir o nível (-) 20 min e a temperatura (Te) assumir o nível (+) 180°C

Situação 2: Quando o tempo (T) assumir o nível (+) 30min e a temperatura (Te) assumir o nível (-) 160°C.

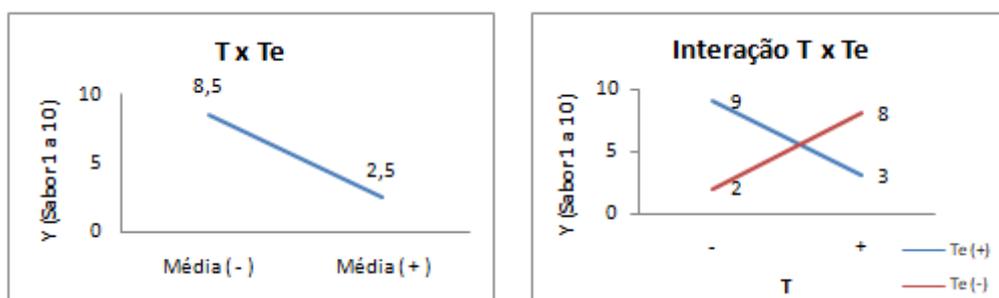


Figura 14 – Gráficos de efeitos da interação T (Tempo) x Te (Temperatura)

Esse é um exemplo hipotético para mostrar 1 ciclo de PDSA, onde tivemos aprendizado sobre a atuação dos efeitos principais e interações sobre a variável resposta, gerando conhecimento suficiente para tomar ação de melhoria. Outras situações poderiam sugerir ciclos subsequentes, caso ainda não fosse atingido o objetivo, dando direcionamento para alteração de níveis ou inclusão de mais fatores.

## 2.7 Aprendizagem de máquina “Machine Learning”

Uma das primeiras abordagens do termo “Machine Learning” foi feita em 1959, por Arthur Samuel em um Journal de pesquisa e desenvolvimento da IBM, onde ele descreve alguns estudos de Aprendizagem de Máquina utilizando jogos de dama. Samuel (1959) declara que já havia trabalho suficiente para entender que um computador pode ser programado para aprender a jogar damas com melhor performance do que a pessoa que escreveu o programa e conclui que esse mesmo conceito de aprendizagem pode ser aplicado a problemas reais.

ML teve seu início no final da década de 1950 e sua presença maior na vida real, décadas mais tarde, principalmente com a evolução da capacidade de processamento de dados dos computadores e a internet.

Vários autores definem ML como o desenvolvimento de habilidades de um computador em realizar tarefas de forma inteligente, alinhados com o pensamento de Arthur Samuel:

“Algoritmos que, mediante uma experiência, melhoram seu desempenho em uma tarefa específica de acordo com uma métrica definida”. (Tom Mitchell, 1997)

“Programas de computador utilizados para otimizar um critério de desempenho, usando dados de exemplo ou experiência do passado” (Alpaydin, 2010)

“Ramo da inteligência artificial que permite que as máquinas executem seus trabalhos com habilidade, usando software inteligente” (Mohammed et al., 2016)

Falar em aprendizagem de máquina envolve vários termos ou disciplinas que auxiliam na integração entre homem e máquina, tais como mineração de dados, estatística, reconhecimento de padrões, base de dados, ciência de dados e inteligência. De uma forma simplificada, o aprendizado de máquina utiliza dados que representam características do mundo real, os quais são processados por algoritmos ou modelos de aprendizado, com capacidade de realizar tarefas. Dentro da mineração de dados, a tarefa de desenvolvimento de um provável futuro (prognóstico) é definida por Côrtes et al.(2002) como:

- *Classificação* - Identifica qual classe/categoria um determinado registro pertence;
- *Estimação ou regressão* - Similar a classificação, porém o registro é identificado por um valor numérico e não categórico;
- *Predição* - Visa prever o valor futuro de um atributo.

Esses modelos passam por avaliações sobre sua habilidade de acertos em relação a previsão.

Mohammed et al. (2016) destacam que o modelo pode ser considerado como um aproximação do processo que queremos que as máquinas imitem. Às vezes, podemos entender o modelo e às vezes não, caracterizando-se como uma caixa preta para nós, onde seu funcionamento não pode ser explicado intuitivamente.

A figura 15 apresenta quatro técnicas de aprendizado de máquina, relacionado aos dados de entrada, com uma breve descrição.

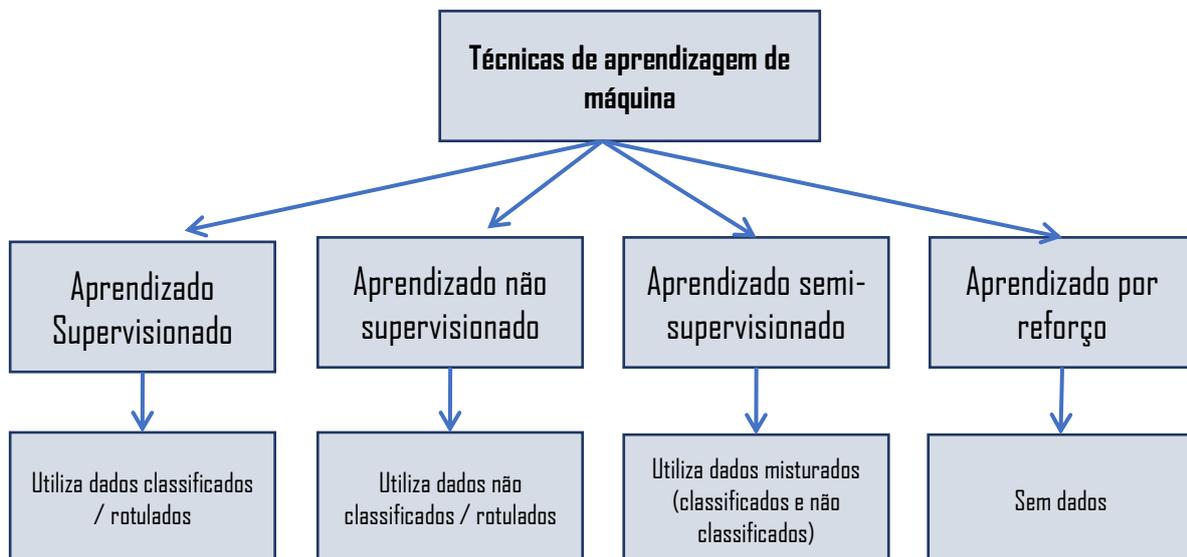


Figura 15 - Técnicas de aprendizado de máquina e seus dados requeridos (Mohammed et al., 2016).

- Aprendizados supervisionados são aqueles nos quais existe um conjunto de dados rotulados com a resposta correta para aprendizagem;

- Aprendizados sem supervisão são aqueles nos quais não existem rótulos para aprendizagem;
- Aprendizado semi-supervisionados são aqueles nos quais apenas alguns dados estão rotulados;
- Aprendizado por reforço é o treinamento de modelos de aprendizado de máquina para tomar uma sequência de decisões em um ambiente potencialmente complexo (Data Science Academy, 2019).

Na sequência, uma breve descrição das técnicas de aprendizado Supervisionado e Não Supervisionado, algoritmos de aprendizado e sua avaliação de performance e o software Orange como plataforma utilizada para projetos de aprendizagem de máquina.

### **Aprendizado Supervisionado**

No aprendizado supervisionado, o objetivo é declarado com dados de treinamento rotulados/classificados. Existe um valor, uma classe atribuída, que permite se ter uma idéia do que está procurando aprender. Mohammed et al. (2016) explicam que os dados de treinamento consistem no vetor de entrada  $X$  e vetor de saída  $Y$  com rótulos. Um rótulo do vetor  $Y$  é a explicação de seus respectivos dados do vetor de entrada  $X$ . Juntos, eles formam um exemplo de treinamento que podem descrever um relacionamento que faça sentido. Essa técnica é chamada de aprendizado supervisionado em função do vetor de saída  $Y$  possuir rótulos/orientação para cada exemplo de treinamento presentes nos dados de treinamento. Mohammed et al., (2016) destacam ainda que os rótulos/orientações para o vetor de saída são fornecidos pelo supervisor, que na maioria das vezes são humanos, mas também podem ser usadas máquinas para essa rotulagem.

Quando criança, pais ensinam seus filhos sobre os nomes (rótulos) de objetos apontando para eles e pronunciando seus nomes, de forma supervisionada.

### **Aprendizado Não Supervisionado**

Segundo Mohammed et al. (2016) no aprendizado não supervisionado não existe um supervisor indicando/rotulando um vetor com um objetivo definido e portanto não existem dados para treinamento. Não há clareza sobre o que se quer aprender. Esse é um fato que os cientistas de dados e profissionais de aprendizado de máquina encontram com frequência, onde há disponibilidade de um volume grande de dados, porém, sem rótulos. Ainda assim, é possível

que se extraíam aprendizados através da busca por estruturas ou comportamentos significantes e ocultos nos dados.

## Algoritmos para aprendizagem de máquina

A seguir é apresentada uma breve descrição sobre os principais algoritmos utilizados em aprendizado de máquina, responsáveis pelo modelo que descreve uma função de relação entre variáveis predictoras (x, atributos) como entradas e variáveis preditivas (Y, classificação, resultado).

### a) Árvores de decisão

Côrtes et al. (2002) definem uma árvore de decisão como um fluxograma (*flow-chart*) semelhante a uma estrutura de árvore, onde cada nó representa um teste em um atributo, cada ramo (sub-árvore) representa o resultado do teste e cada folha representa a distribuição dos registros. São os modelos mais usados em inferência indutiva, onde o modelo é treinado de acordo com um conjunto de exemplos previamente classificados, que o caracteriza como aprendizado supervisionado, permitindo a classificação ou previsão futura com base na árvore modelada.

Na figura 16, Côrtes et al. (2002) nos fornecem um exemplo de aplicação de uma árvore de decisão, onde os dados de treinos foram classificados com “Comprador” e “Não Comprador”. Observa-se que 90% dos homens com salário superior a R\$4.000,00 são candidatos a comprarem o produto.

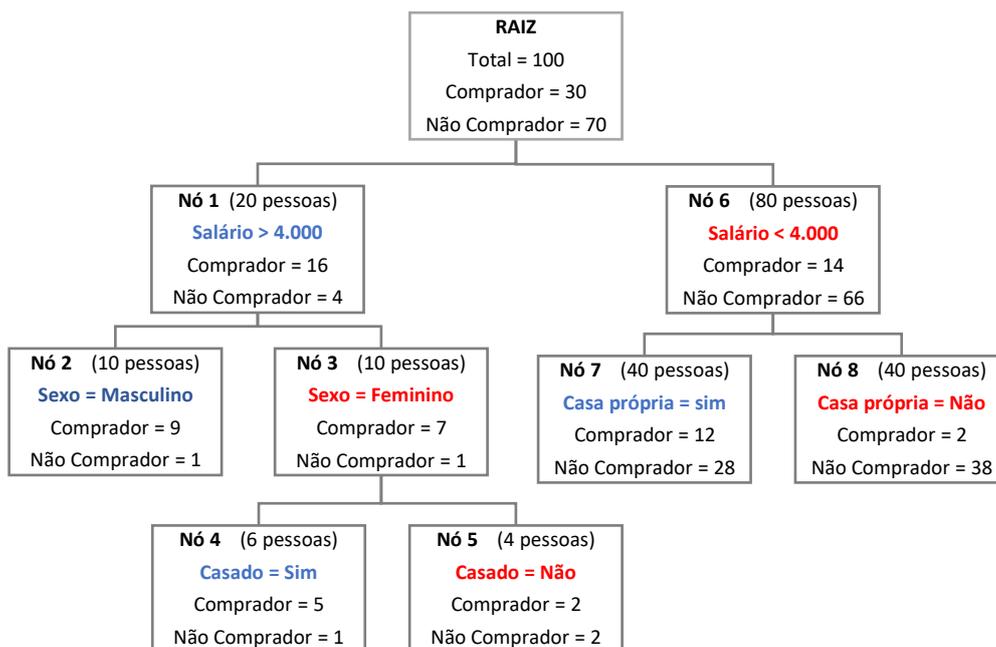


Figura 16 – Exemplo de visualização de uma árvore de decisão – Adaptada de Côrtes et al. (2002).

### b) Máquina de Vetores de Suporte (Support Vector Machine)

É um algoritmo supervisionado utilizado para reconhecimento de padrões e tarefa de classificação. Segundo Côrtes et al. (2002), a separação de classes é feita por um hiperplano que utiliza vetores de suporte (conjunto de treinamento), tornando visível a separação de grupos. A figura 17, mostra o exemplo de dados de treinamentos classificados como “abaixo da meta” (bolinhas cinzas) e “acima da meta” (bolinhas brancas), onde o algoritmo busca identificar grupos pela presença dos hiperplanos que separam as classes.

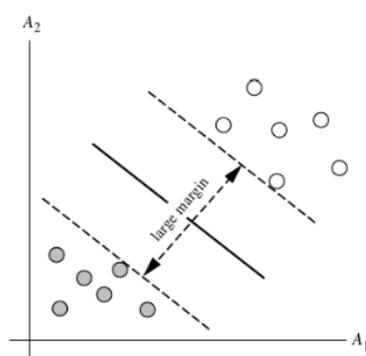


Figura 17 - Dados de treinamento, onde existe um número infinito de hiperplanos que podem separar as classes (Côrtes et al., 2002).

### c) Regressão Logística

A regressão logística é um dos algoritmos para aprendizado de máquina usado para classificação, onde a variável de resposta é categórica.

Em um resumo feito por (Lin, 2018), o modelo funciona prevendo a probabilidade de  $Y$  pertencer a uma categoria específica, primeiro ajustando os dados a um modelo de regressão linear, que é então passado para a função logística. A função logística sempre produzirá uma curva em forma de S, portanto, independentemente de  $X$ , pode-se obter sempre uma resposta sensata (entre 0 e 1). Se a probabilidade estiver acima de um certo limite predeterminado (por exemplo,  $P(\text{Sim}) > 0,5$ ), então o modelo irá prever Sim.

### d) Naive Bayes

Trata-se de um método de aprendizagem supervisionado, utilizando algoritmos baseados na aplicação do teorema de Bayes. Camilo & Silva (2009) descrevem o teorema de Bayes, como sendo possível encontrar a probabilidade de um certo evento ocorrer, dada a probabilidade de

um outro evento que já ocorreu: Probabilidade (B dado A) = Probabilidade (A e B) / Probabilidade(A).

O algoritmo de naive Bayes leva em consideração que não existe relação de dependência entre os atributos.

### Avaliação da performance de algoritmos – Matriz de confusão

Para avaliar a performance do algoritmo/modelo, alguns indicadores são considerados relacionados a sua habilidade de prognóstico utilizando os dados de treino e teste: Estimar, prever ou classificar as novas instâncias.

A matriz de confusão mostra as classificações previstas e reais de um algoritmo de aprendizado de máquina (Kohavi, R; Provost, 1998). Visualmente fornece os acertos e erros do algoritmo, de acordo com seu impacto nos processos em que estão inseridos – processos de decisões de manufatura ou diagnóstico de doenças, por exemplo.

Na tabela 5, está representada a matriz de confusão de um algoritmo de classificação binária (sim/não, doente/sadio, defeituoso/não defeituoso, positivo/negativo, pertence/não pertence), onde a previsão do algoritmo é confrontado com o resultado real/verdadeiro e conhecido. Essa matriz dispõe de 4 possibilidades de classificação do algoritmo.

Tabela 5– Matriz de confusão de classificação binária (Kohavi, R; Provost, 1998)

		Previsão do Algoritmo	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

- Verdadeiro Positivo (VP) – Algoritmo previu “Sim” & Resultado real “Sim”
- Verdadeiro Negativo (VN) – Algoritmo previu “Não” & Resultado real “Não”
- Falso Positivo (FP) – Algoritmo previu “Sim” & Resultado real “Não”
- Falso Negativo (FN) – Algoritmo previu “Não” & Resultado real “Sim”

Avaliando a matriz de confusão, tabela 5, os acertos do algoritmo encontram-se na diagonal azul (VP e VN) e os erros encontram-se fora da diagonal azul (FP e FN).

Com a matriz de confusão, é possível obter as métricas típicas utilizadas para avaliar algoritmos de ML. Os significados e fórmulas das métricas são apresentadas abaixo de forma sucinta, com base em Matos et al. (2009):

- **CA (Classification Accuracy)** - Mede a habilidade do algoritmo quanto ao acerto de suas previsões, ou seja, mede a fração de previsões corretas (VP e VN). Essa métrica não é apropriada quando a classificação das instâncias para treinamento são desproporcionais e muitas informações são irrelevantes.

$$CA = \frac{(VP+VN)}{(VP+VN+FP+FN)} \quad (1)$$

- **Precision** - Taxa que avalia a capacidade do algoritmo classificar/prever como positivos corretamente, ou seja, mede a fração de positivos reais em relação a todos os casos que são previstos como positivos.

$$Precision = \frac{(VP)}{(VP+FP)} \quad (2)$$

- **Recall**- Taxa que avalia a capacidade do algoritmo classificar/prever como positivos aquilo que realmente é positivo, ou seja, mede quantos positivos reais são previstos como positivos.

$$Recall = \frac{(VP)}{(VP+FN)} \quad (3)$$

- **Cálculo de F1**- É a média harmônica entre precisão e recuperação/Recall

$$F1 = \frac{(2 \times Precisão \times Recall)}{(Precisão + Recall)} \quad (4)$$

Os diversos algoritmos de aprendizado de máquina são submetidos as avaliações de AC, Precision, Recall e F1, onde as melhores performances direcionam para a escolha do melhor algoritmo.

## 2.8 Software “Orange”

Orange é um programa, de código aberto, que trabalha com visualização de dados e aprendizado de máquina. Possui uma interface amigável intuitiva, permitindo a construção de um fluxo de trabalho com base em análise de dados e uma ampla opção de ferramentas, considerando técnicas de visualização, exploração, estatística, pré-processamento e modelagem de dados (LARHUD, 2018).

O software foi desenvolvido pelo [\[Laboratório de Bioinformática\]](#) dentro da Faculdade de Informação e Computação na Universidade de Ljubljana localizada na Eslovênia, tem mais de 20 anos e encontra-se na versão 3.27.

O programa trabalha com arquivos nos formatos abaixo:

- Excel, arquivos delimitados por vírgulas e tabulações (.xlsx, .csv, .tab);

- Dados on-line no format do Google Planilhas;
- Imagens (.jpg, .tiff, .png);
- Arquivos de texto (.txt, .docx, .odt);
- Banco de dados PostgreSQL e MSSQL.

Com uma base de dados, conforme fomatos acima, é possível criar um fluxo, utilizando as ferramentas disponíveis para execução do estudo. A figura 18 mostra como o Software Orange se parece, com uma área para a construção do fluxo de análise de dados, onde as ferramentas, representadas por ícones, são extraídas dos grupos de ferramentas que se encontram em um painel principal, do lado esquerdo da tela, bastando arrastar os ícones e trazendo para o Canvas.

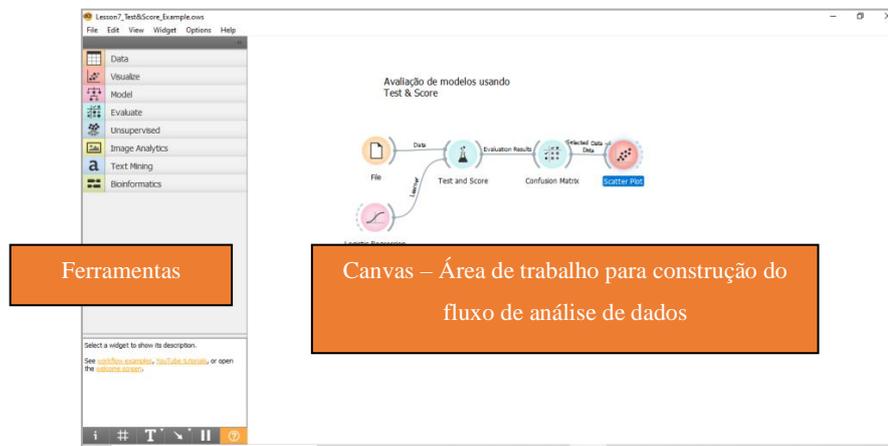


Figura 18 – Visualização do programa Orange



Figura 19 – Conjunto de ferramentas de processamento

A figura 19 mostra os principais conjuntos de ferramentas disponíveis para utilização:

- **[Data]**: manipulação de dados;
- **[Visualize]**: visualização de dados;
- **[Model]**: predições;
- **[Evaluate]**: classificação avaliativa e desempenho de algoritmos;
- **[Unsupervised]**: processamentos mais customizados.

Para aprender como usar o Orange, o software disponibiliza vídeos com tutoriais no youtube, bem como exemplos prontos, facilitando o aprendizado principalmente para iniciantes, conforme mostra a figura 20.

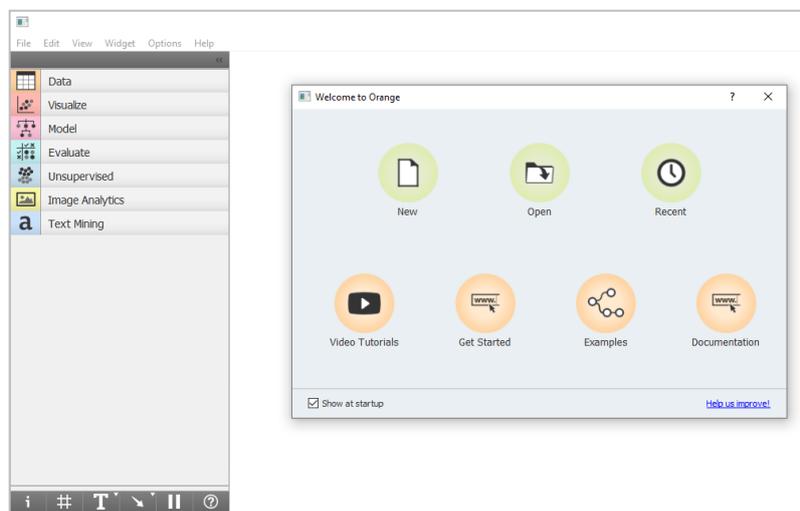


Figura 20 – Tela inicial do programa Orange– Fonte: Software Orange

### 3 Definição do problema e proposta de solução

#### 3.1 Definição do problema

Abaixo, temos o fluxo simplificado de um processo de fabricação ar condicionado de janela, de uma empresa do polo industrial de Manaus, com etapas de montagens e testes/inspeções para garantir a qualidade do produto final.

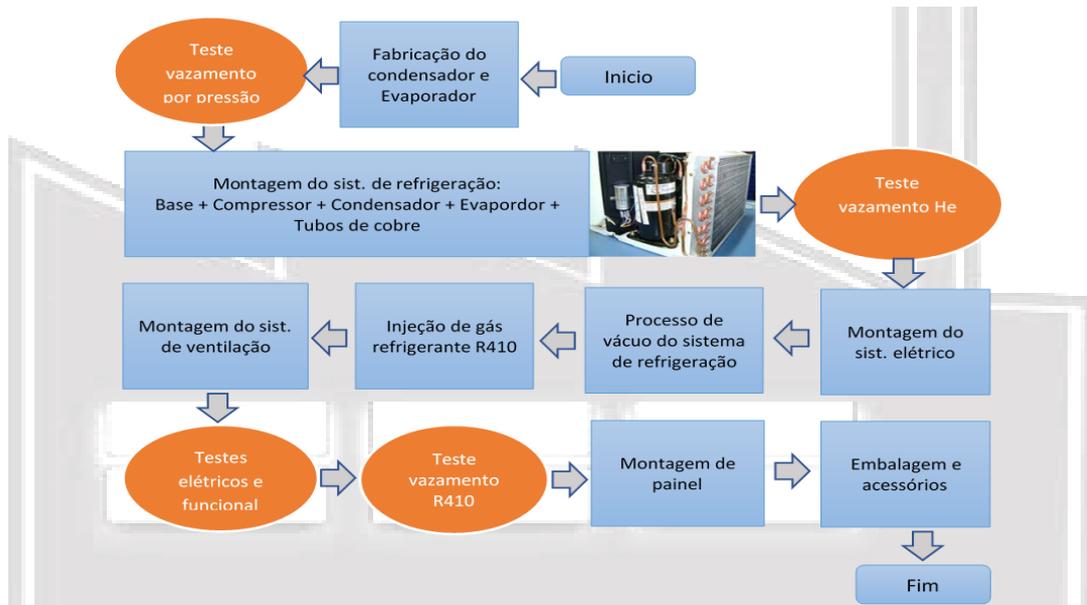


Figura 21 – Processo de montagem de ar condicionado de janela

Os principais testes e inspeções, realizados em 100% dos produtos, na linha de montagem de ar condicionado estão descritos abaixo:

- 1 *Teste de vazamento por pressão* positiva em trocadores de calor - Aplicado em trocadores de calor (Evaporador/Condensador), duas peças que fazem parte do sistema de refrigeração, compostas de tubo de cobre e alumínio. Essa peça contém uma quantidade grande de junções de tubo de cobre, feitas através da brasagem de solda, o que aumenta a probabilidade de vazamento de gás nas junções. O teste consiste em aplicar ar comprimido no interior dos trocadores de calor a uma pressão positiva, observando se existe a presença de queda de pressão, o que indica vazamento
- 2 *Teste de vazamento de He* - Aplicado ao sistema selado completo. Utiliza o método de espectrometria de massa com gás de teste He. No interior do sistema o gás é injetado a uma pressão positiva e um equipamento detector de presença de He, também conhecido como Sniffer, é utilizado por um operador para identificar a presença de gás na parte externa ao produto. O gás He é utilizado apenas para o momento do teste,

onde existe no processo um sistema de injeção do gás, o teste é realizado e depois o gás é recolhido.

- 3 *Testes elétricos e funcionais* - Realizados para garantir a segurança elétrica e o funcionamento adequado do produto;
- 4 *Testes de vazamento R410*- Aplicado ao sistema selado completo, utilizando o método de espectrometria de infravermelho com gás refrigerante. O gás refrigerante R410 é injetado no interior do sistema a uma pressão positiva e um equipamento detector de gás R410 é utilizado por um operador para identificar a presença de gás na parte externa do produto. O R410 é o gás refrigerante que faz parte da estrutura do produto e caso aconteça a fuga desse gás, o produto perde sua função principal de refrigerar o ambiente.

Uma das etapas mais críticas desse processo para a qualidade do produto é a garantia do sistema de refrigeração selado, sem furos que permitam vazamento de gás refrigerante. Várias junções entre tubos de cobre fazem parte da estrutura do produto, necessárias para fechar o sistema por onde o gás circula, conforme ilustrado na figura 22. Essas junções são feitas através de uma liga depositada nas junções, por meio do processo de brasagem. Essas são as posições mais vulneráveis de apresentar o vazamento de gás.

A ausência de gás refrigerante dentro do sistema de refrigeração causa a falha da principal função do ar condicionado, a de refrigerar ou aquecer o ambiente.

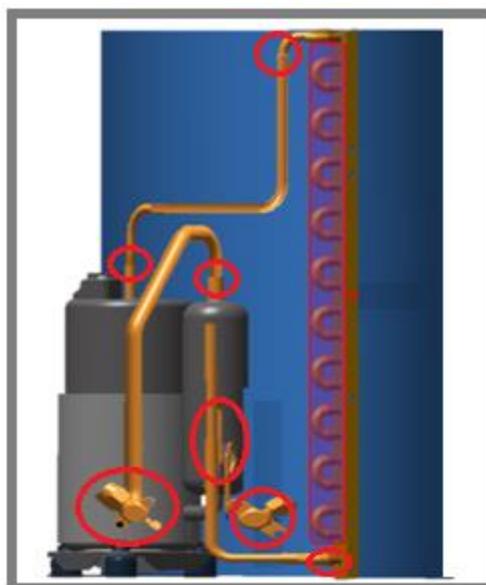


Figura 22 - Pontos de junções de solda em um trocador de calor

No fluxo de processo atual existem dois testes de vazamento de gás, utilizando métodos de alta pressão no sistema de refrigeração, variando apenas o tipo de gás. A efetividade desse teste depende da atuação entre equipamentos de medição, produto em teste e operador, que tem uma participação importante para garantir o método de teste (aproximação dos pontos de solda e velocidade da ponta de prova o equipamento).

Mesmo com todos os testes atuais no processo, ainda existem produtos que chegam a casa do consumidor sem o gás refrigerante ou com gás insuficiente. Essa afirmação pode ser identificada em resultados de um teste de vazamento implementado no depósito de produto acabado, nas dependências da fábrica, onde um operador aplica a ponta de prova de um detector de gás refrigerante, através de um furo na embalagem, identificando a presença ou não do gás refrigerante. Esses produtos voltam para a fábrica para reparo e identificação de onde o vazamento acontece, com foco em melhoria contínua. A figura 23 apresenta a quantidade de produtos com falha de vazamento.



Figura 23 – Gráfico de produtos com vazamento no depósito

O proceso atual é uma evolução de várias melhorias, com foco na não geração de vazamento, bem como na detecção/contenção de vazamento de gás.

### 3.2 Proposta de solução

Esse estudo busca alternativas que melhore a detecção, reduzindo a interferência humana no julgamento, atendendo a velocidade de produção industrial, com baixa ocupação de área.

A alternativa a ser estudada é a utilização de imagens com radiação infravermelho (IR) na detecção do vazamento de gás durante o processo industrial do ar condicionado. Uma forma de enxergar o gás que é imperceptível aos sentidos humanos.

Um teste inicial foi feito, de forma empírica com uma câmera IR, marca FLIR, utilizada pela área de manutenção para diagnóstico de falhas em quadro elétrico. Foi simulando um vazamento nas válvulas de serviço do aparelho condicionador de ar e foi possível identificar a presença de uma nuvem de gás ao redor do vazamento, o que não é possível com câmera convencional, como mostra a figura 24.

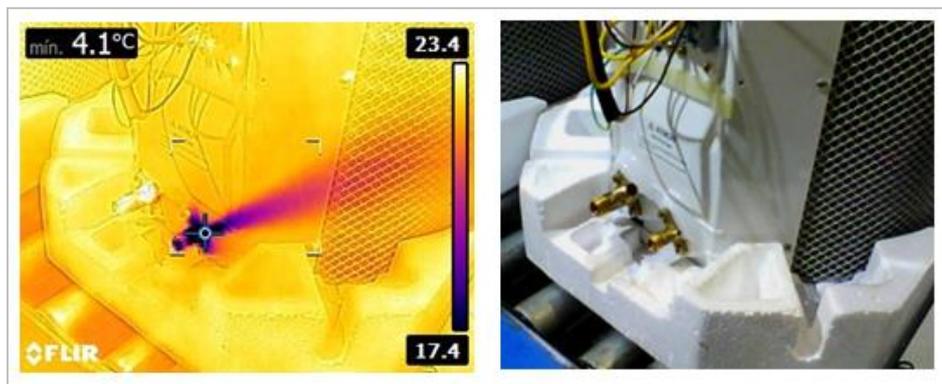
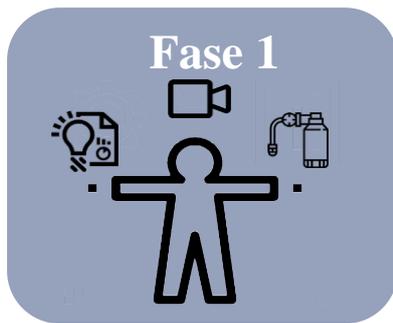


Figura 24 – Imagem de um produto com vazamento de gás refrigerante (Lado esquerdo com câmera de IR e Lado direito com câmera convencional)

Algumas perguntas foram feitas para direcionar o estudo:

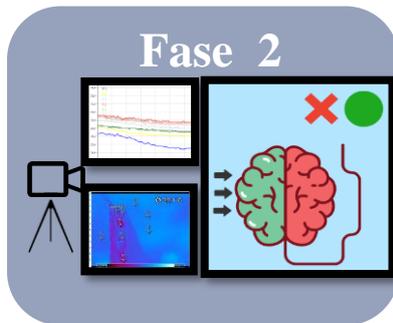
- É possível detectar vazamento de gás utilizando camera de IR de até 3 gramas/ano (referência do teste atual com equipamento que detecta He)?
- É possível detectar vazamentos em até 20 segundos (tempo máx de operação em um posto de teste do processo em avaliação)?
- Quais os fatores que podem influenciar no resultado do teste?
- É possível utilizar os dados da câmera para aprendizagem de máquina com classificação de um produto com vazamento?

Para responder as perguntas acima, serão desenvolvidos as fases 1 e 2, conforme figura 25, para testar o conceito da proposta de usar uma câmera de infravermelho na detecção de vazamento. Técnicas de experimentos sequenciais foram utilizadas, onde vários fatores (x) de processo e produto foram considerados para análise dos seus efeitos na variável resposta (Y), bem como a utilização dos dados coletados pela câmera para aplicação em aprendizagem de máquina, a fim de classificar a presença de vazamento. As fases 3 e 4 são etapas futuras desse projeto.



### ***Estudo Industrial–Teste de Conceito***

- Gás (tipo, Temperatura, Pressão)
- Tamanho de vazamento
- Câmera IR (distância, resolução, ...)
- Ruídos
- Tempo para detecção do vazamento



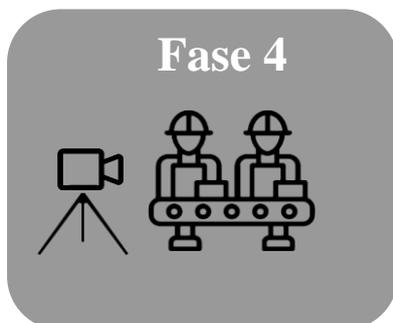
### ***Aprendizagem de máquina***

- Tipos de dados da câmera
- Dados para treino e teste
- Teste de performance de algoritmos
- Algoritmo de classificação de vazamento



### ***Projeto do teste***

- Câmera – (quantidade)
- Posição da câmera
- Automação
- Desenvolvimento da Inteligência de decisão
- Protocolo de comunicação com MES



### ***Integração Industrial***

- Posição do teste no processo
- Instalação
- Integração de TI
- Processo de operação do teste
- Lote piloto Produção com novo teste
- ...

Figura 25 - Fases de desenvolvimento do projeto de teste de vazamento

## 4 Teste do conceito – PDSA's

Nesse capítulo será descrito a metodologia de planejamento, coleta e análise de dados de experimentos sequenciais utilizando ciclos de PDSAs, que vão evoluindo a medida que o conhecimento vai sendo adquirido, buscando sempre uma relação de causa e efeito  $Y=f(x)$  e o teste do conceito para uso de câmeras de IR para detecção de vazamento.

### Mapeamento das variáveis envolvidas no estudo

Para avaliar o impacto de fatores que possam provocar variações na variável resposta, foi feito um brainstorming e Ishikawa com os conhecedores do processo e produto e encontram-se abaixo:

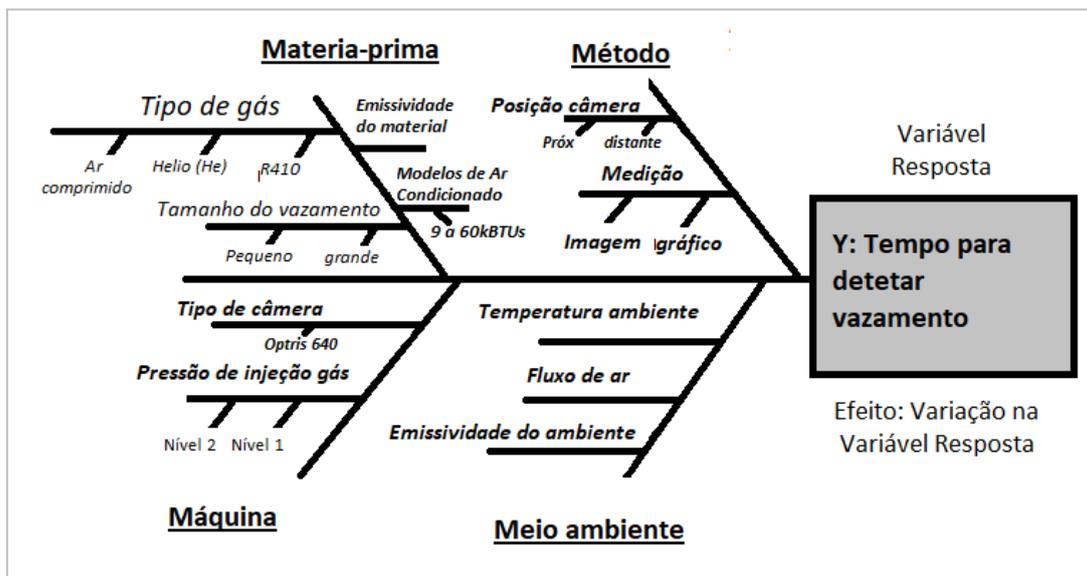


Figura 26 – Diagrama de causa e efeito: levantamento de variáveis

### Variável resposta (Y):

(Y1) Tempo que leva para detectar uma redução na temperatura de  $0,5^{\circ}\text{C}$

(Y2) Tempo para visualizar o vazamento

Requisito: A detecção deve acontecer dentro de 20 segundos

### Variáveis causais ou independentes – fatores principais

(A) *Tamanho e tipo do Vazamento* – São variados os tipos, formatos e tamanhos de vazamento.

Os tipos de vazamento mais comuns encontrados na fabricação de ar condicionados, acontecem principalmente em junções de tubos de cobre, conexões permanentes unidas por um material de adição, que derrete no processo de aquecimento com chama e solidifica-se proporcionando a conexão permanente, ilustrados na figura 27.

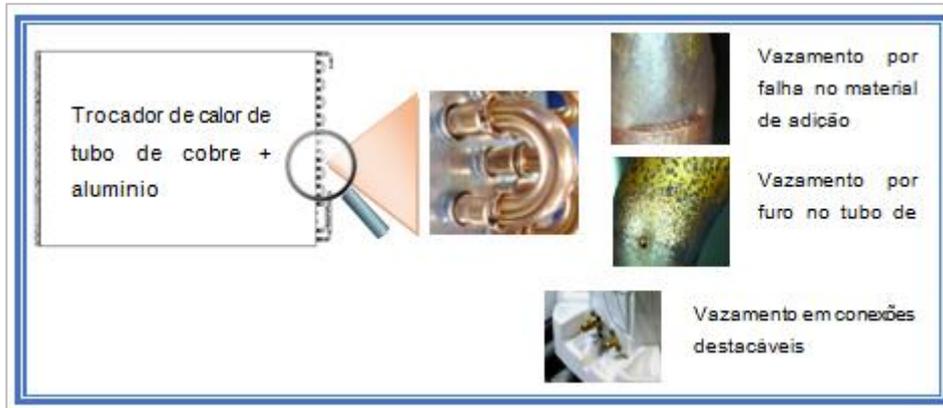


Figura 27 – Tipos e tamanhos de furos que permitem o vazamento de gás

(B) *Pressão do gás He* – Para que o gás seja inserido é necessário uma pressão de inserção maior que a pressão atmosférica. A pressão de inserção do gás no processo atual é de 6 Bar

(C) *Distância entre posição da câmera e produto*

(D) *Tipo de gás* – O processo atual trabalha com 3 tipos de gases:

- ar comprimido, utilizado em processos pneumáticos e testes de vazamento em baixa pressão
- Helio (He).
- Refrigerante 410a

(E) *Tipo de camera* – A camera em estudo é a Optris PI640, a qual não temos histórico de uso em processos similares para detecção de vazamento.

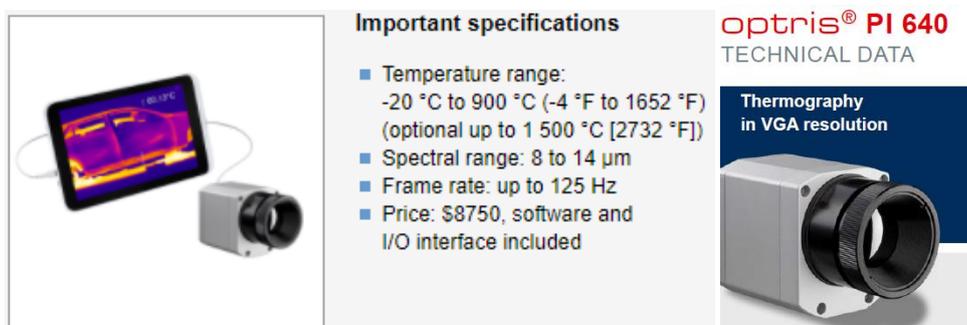


Figura 28 – Camera Optris PI640 – Fonte: <https://www.optris.com/thermal-imager-optris-pi-640>

(F) *Modelo de produto a ser testado* – A empresa possui diferentes capacidades de produto, variando de 9.000BTUs 60.000BTUs. Será considerado inicialmente o modelo 12.000BTUs, por estar entre os de maior volume de produção.

#### 4.1 PDSA#1 “Plan Do Study Action”

##### Objetivo:

Avaliar fatores de processo e método de medição com a camera de IR para identificar vazamento de gás em um produto de ar condicionado com possibilidades de aplicação em processo de fabricação em larga escala, dentro do tempo de 20 segundos.

##### Informações e conhecimento prévio adquiridos sobre o assunto

Tem se o conhecimento de que câmeras térmicas infravermelho captura diferença de temperatura na superfície de um objeto. Uma utilização comum na indústria é a identificação de áreas com temperatura elevada em quadros elétricos. Para o estudo sobre vazamento de gás, não foi encontrado na literatura aplicações do tipo.

##### Estratégia de experimentação

Neste experimento foram estudados inicialmente 3 fatores (A, B e C) em 2 níveis diferentes para cada fator (tabela 6). Alguns fatores identificados inicialmente foram considerados constantes durante o experimento (tipo de gás He, Tipo de câmera, Modelo de ar condicionado), bem como o cuidado com as variáveis de ruído, conforme tabela 7.

Tabela 6 - Fatores e níveis – DOE#1

Fatores	Nível (-)	Nível (+)	Teoria
A – Tamanho e posição do furo	Furo grande, posição externa, na curva (0,4mm <sup>2</sup> )	Furo pequeno, posição externa, na curva (0,3mm <sup>2</sup> )	Quanto menor o tamanho do furo, mais tempo levará para detectar o vazamento
B – Pressão de injeção do gás	3,5bars	5bars	Quanto menor a pressão do gás, mais tempo levará para detectar o vazamento
C – Distância da câmera para o produto	15cm	50 cm	Quanto maior a distância, mais tempo levará para detectar o vazamento
D – Tipo de gás	He		Durante esse experimento o gás utilizado será somente o He
E – Tipo de câmera	Modelo Optris 640		Durante esse experimento será utilizado apenas um modelo de câmera
F – Modelo de ar-condicionado	12.000 BTUs		Durante esse experimento será utilizado apenas um modelo de ar-condicionado. Pertence à família de maior volume

Tabela 7 - Variáveis de ruído – DOE#1

Variáveis de ruídos	Métodos de controle
Fluxo de ar ambiente	Controlar pessoas se movimentando na área do experimento
Temperatura ambiente	Monitoramento
Setups dos níveis dos fatores	Monitoramento/Observar restrições

A estratégia também pode ser representada por uma árvore de amostragem, conforme figura 29. O experimento foi conduzido em laboratório.

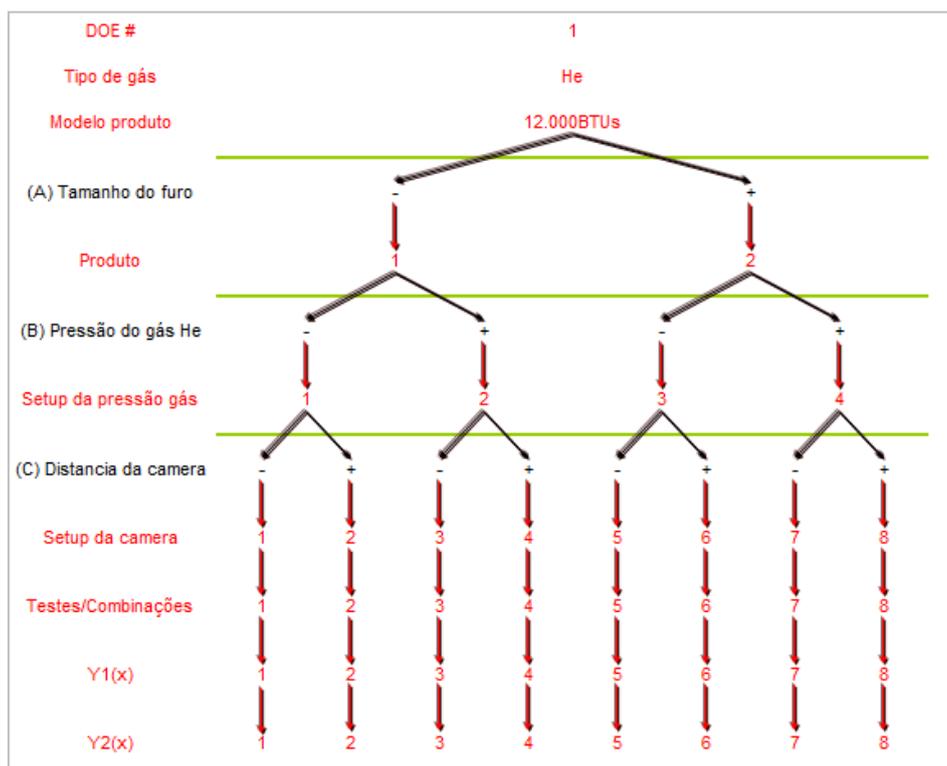


Figura 29 - Árvore de amostragem do experimento – DOE#1

### Matriz de experimentação, coleta e análise dos dados - DOE#1

A matriz de experimentação apresentada na tabela 8, possui 8 combinações (testes) diferentes, variando os fatores A (tamanho do furo), B (Pressão de injeção de gás) e C (Distância da camera), conforme seus níveis (-1 e +1) que na prática correspondem aos níveis descritos na parte superior da tabela. Essa é a parte da matriz de experimentos que é manipulada durante o experimento e os valores da variável resposta, nas colunas Y1 e Y2 são observados e coletados.

Tabela 8 - Matriz de experimentação e coleta de dados – DOE#1

Níveis									
(-1)	Furo grande	3,5 Bar	15cm	Interações					
(+1)	Furo pequeno	5 Bar	50cm						
Testes	(A) Tamanho do vazamento	(B) Pressão de injeção do gás	(C) Distância da câmara	AB	AC	BC	ABC	Y 1 (Tempo p reduzir 0,5°C -seg)	Y2 (Tempo para visualizar o vazamento - seg)
1	-1	-1	-1	1	1	1	-1	15	12,3
2	-1	-1	1	1	-1	-1	1	9,5	9,8
3	-1	1	-1	-1	1	-1	1	9	8
4	-1	1	1	-1	-1	1	-1	8	8
5	1	-1	-1	-1	-1	1	1	20	14
6	1	-1	1	-1	1	-1	-1	18	15
7	1	1	-1	1	-1	-1	-1	6	20
8	1	1	1	1	1	1	1	19	12
Efeito (Y1)	5,4	-5,1	1,1	-1,4	4,4	4,9	2,6		
Efeito (Y2)	5,7	-0,8	-2,4	2,3	-1,1	-1,6	-2,9		

As colunas de interações são construídas a partir das colunas dos fatores principais (A, B e C) e portanto, não são manipuláveis. No entanto, com esse experimento é possível estudar sobre o efeito dos Fatores Principais e as interações possíveis entre eles (AB, AC, BC e ABC) e a magnitude dos efeitos pode ser calculada com base nos dados da variável resposta Y e observada na parte inferior da tabela.

### Medição da variável resposta (Y1 e Y2)

Tomando como exemplo o teste 1, foi montado um produto com as seguintes condições:

A (-): Vazamento com furo de tamanho 0,4mm<sup>2</sup>

B (-): Pressão de injeção do gás He em 3,5bar

C (-): Distância da camera para o ponto de vazamento em 15cm

O produto foi posto em funcionamento e observado com a camera por 20 segundos, com possibilidades de medições de temperatura na região com vazamento e regiões sem vazamento, capturando imagem térmica e gráfico linear de temperatura versus tempo.

As figuras 30 e 31 mostram o resultado do teste 1, que fornece imagem térmica do objeto (lado esquerdo) e o valor de temperatura ao longo do tempo da área do vazamento (lado direito).

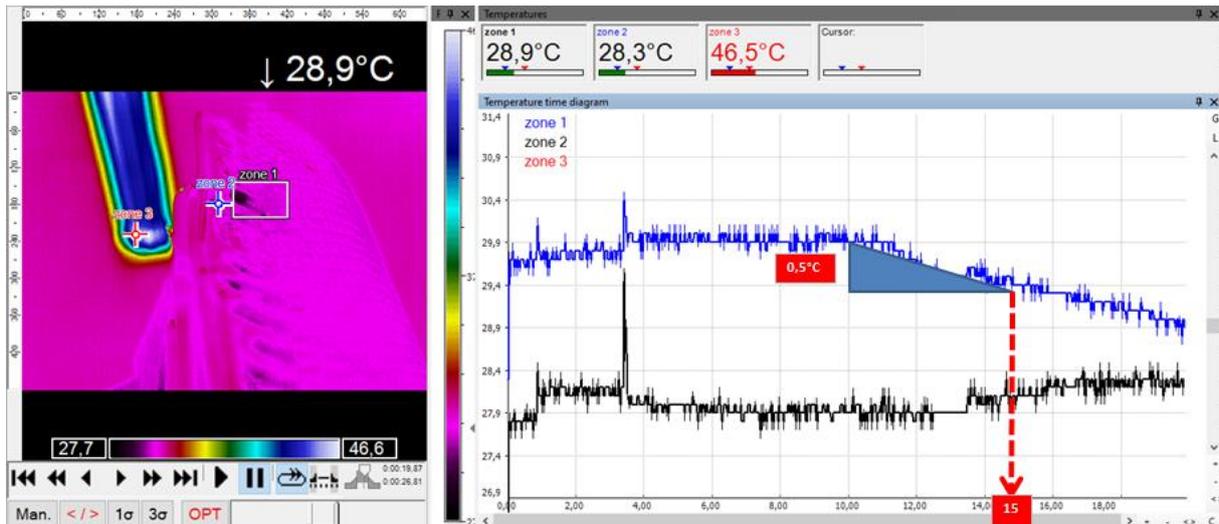


Figura 30 - Medição da variável resposta Y1 (tempo para reduzir a temperature em 0,5°C na região do vazamento: 15 segundos)

Uma característica observada na região de vazamento é a queda de temperatura ao longo do tempo e que uma queda de 0,5°C ou mais caracteriza-se a presença de um vazamento. A figura 30 mostra a medição do Y1, onde é observado o tempo em que ocorre uma queda 0,5°C de temperatura na região com vazamento.

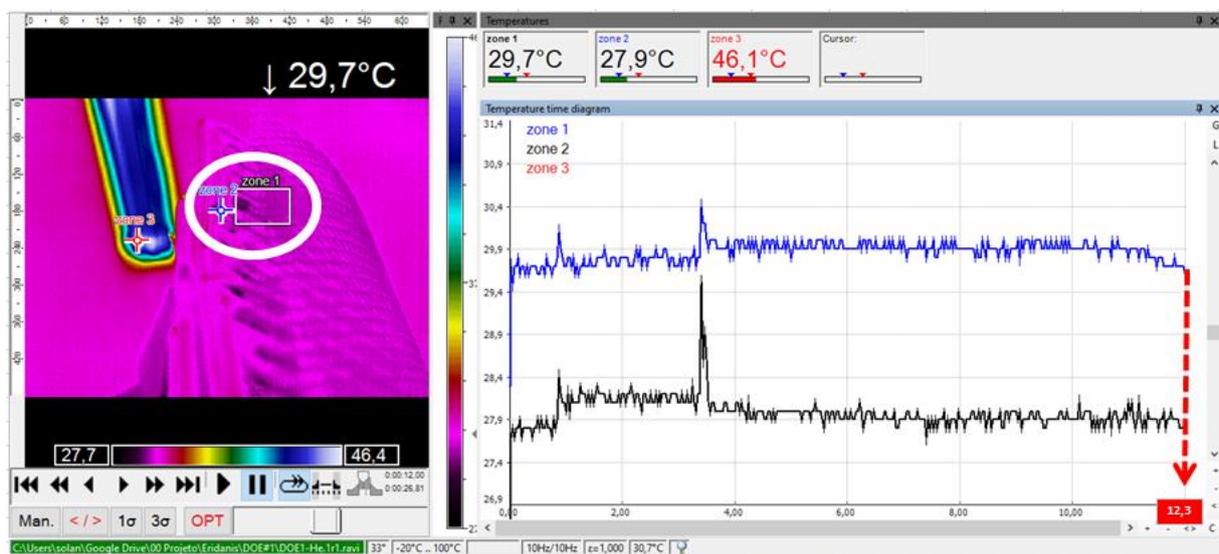


Figura 31 - Medição da variável resposta Y2 (tempo para visualizar a imagem do vazamento, seja gás escapando ou mudança de cor permanente: 12,3segundos)

A figura 31 mostra o resultado de Y2 para o teste 1, onde foi observado o tempo em que acontece uma mudança de cor permanente na região do vazamento, indicando uma temperatura mais fria. Nesse caso, aconteceu com 12,3 segundos.

Para todos os testes foi adotado esse método de medição e os valores encontram-se nas duas últimas colunas da tabela 8.

### Análise prática

Para análise prática, com auxílio gráfico na figura 32, tem-se o resultado do DOE#1 apresentando os 8 testes na sequência, onde se obteve uma variação de 8 segundos a 20 segundos. O valor de 20s foi assumido para valores maiores ou igual a 20 segundos, visto que o teste foi limitado ao tempo máximo de 20 segundos.

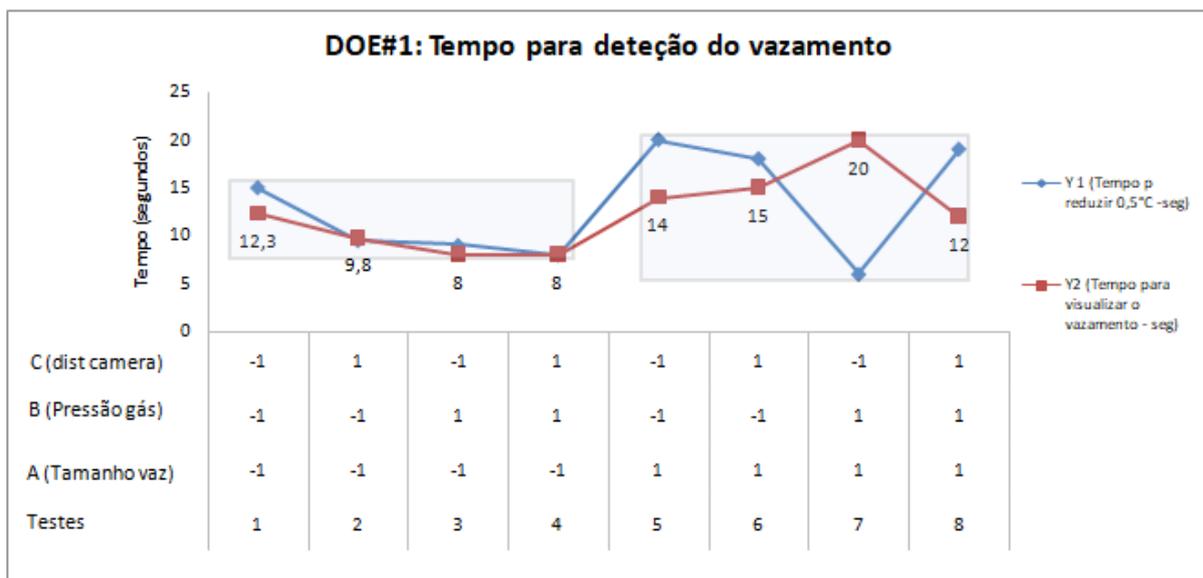


Figura 32 – Tempo para detecção do vazamento Y1 e Y2 – DOE#1

Os 4 primeiros testes (1 ao 4) apresentaram valores entre 8 e 15 segundos e os dois Y's praticamente concordam entre si. Os 4 testes finais (5 ao 8) apresentaram uma discordancia maior entre os Y's e apresentaram variação de 7 a 20 ou acima de 20 segundos. Ou seja, os 4 primeiros testes, que estão relacionados ao fator A (-) tamanho do furo maior tem uma consistencia maior no resultado e leva menos tempo para identificar o vazamento. Quando usado o fator A (+) furo menor leva mais tempo para detetar o vazamento, as vezes passando do tempo limite da linha de montagem.

## Análise gráfica

Para análise gráfica, a figura 33 mostra a magnitude dos efeitos principais e interações que impactam na variável resposta Y1 e Y2.

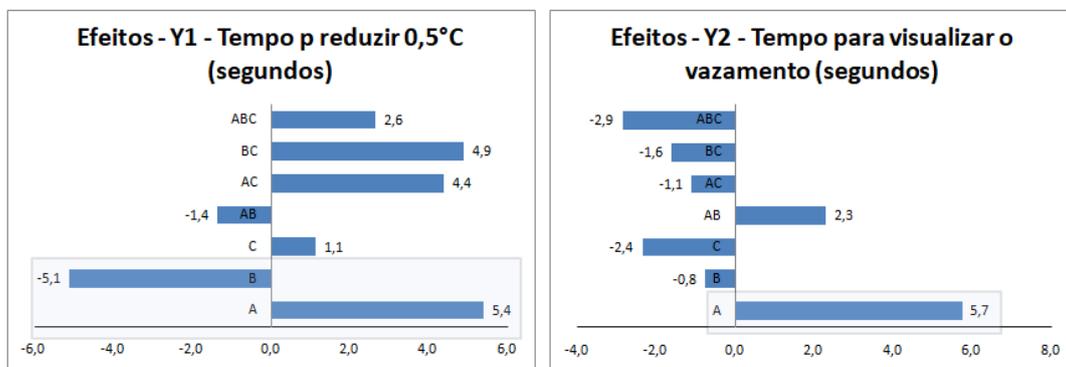


Figura 33 – Gráfico de Efeitos Principais e Interações – DOE#1

O resultado positivo ou negativo no gráfico, indica aumento ou redução na variável resposta quando o fator muda do nível (-1) para o nível (+1). Por exemplo, fator A provoca um aumento de 5.4 segundos quando muda do (furo grande) para (furo pequeno), enquanto que o fator B provoca uma redução de 5.1 segundos quando muda de (3,5Bar) para (5Bar).

Observando os dois Ys, o fator A (tamanho do furo) é o que produz maior impacto. O segundo maior impacto vem do fator B (pressão de injeção do gás) somente para o Y1.

Para um experimento inicial exploratório, há um maior interesse em fatores principais e menos em interação entre eles, em função de ruídos presentes no experimento.

A mesma informação pode ser visualizada nos gráficos de efeitos principais (figuras 34 e 35) e gráfico de interações (figura 36).



Figura 34 - Gráficos de efeito principal para Y1 (tempo para reduzir 0,5°C)

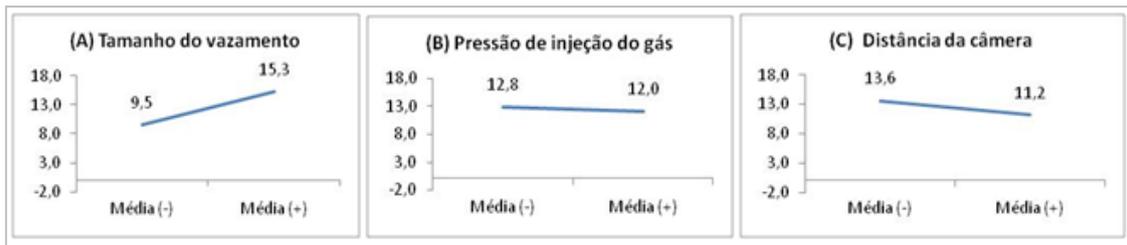


Figura 35 - Gráficos de efeito principal para Y2 (tempo para visualizar a imagem do vazamento)

Fator A: O produto com tamanho de vazamento menor (A+) leva mais tempo para ser detectado, nos dois Y's;

Fator B, o produto com pressão maior (B+) leva menos tempo para ser detectado para o Y1 e não significativo para o Y2;

O fator C não apresentou variação significativa com a posição da camera próxima ou distante do ponto de vazamento para os dois Y's

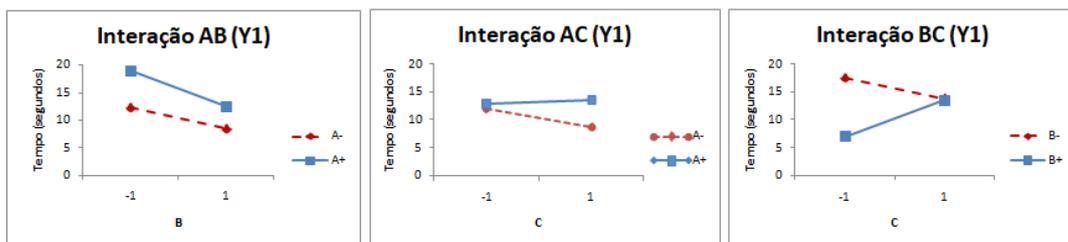


Figura 36 - Gráficos de interações de segunda ordem para Y1 (tempo para reduzir 0,5°C)

Avaliando as interações AB, AC e BC, a figura 30 (efeitos), mostra uma magnitude elevada para AC e BC. Pode-se confirmar nos gráficos de interação, figura 36, que existe uma interação mais forte entre AC e BC. Para a interação AB, esta se mostrou fraca.

## Aprendizados e próximos passos

Com esse primeiro experimento, foi possível aprender sobre os seguintes pontos:

- 1) Comportamento visual da temperatura na região do vazamento, como forma de medição. Dois resultados podem ser capturados ao mesmo tempo: a) Imagem termográfica, onde a região do furo se torna uma região escura identificando uma região mais fria e fica constante enquanto o gás permanece vazando; b) A temperatura pode ser capturada graficamente na região do vazamento e uma queda de temperatura pode ser observada. Utilizando essa forma de medição, foi possível identificar fatores

ativos para detetar o vazamento através da imagem ou queda de temperatura, gerando variação dentro do limite de operação da linha de produção de 20 segundos.

- 2) O fator A (tamanho do vazamento) se mostrou significativo, levando um tempo maior para ser detectado quando o furo é menor.
- 3) O fator B (pressão de inserção do gás He) se mostrou como o segundo mais significativo, provocando uma redução no tempo de detecção quando a pressão é mais elevada.
- 4) Interação BC (Pressão do gás e Distância da câmera) indica tempo de detecção menor, quando B (+1 pressão maior) e C (-1 Distância câmera menor)

Para os próximos experimentos, os mesmos fatores serão considerados, alterando apenas os níveis com o objetivo de se obter a detecção com o menor tempo possível e aumentando a confiança nos resultados.

## **4.2 PDSA #2**

### **Objetivo:**

Avançar nos estudos para avaliar fatores de processo e método de medição com a camera de IR para identificar vazamento de gás em um produto de ar condicionado com possibilidades de aplicação em processo de fabricação em larga escala, dentro do tempo de 20 segundos.

### **Informações e conhecimento prévio adquiridos sobre o assunto**

Foi possível identificar fatores ativos (A - Tamanho do furo, B – pressão de injeção He) e uma possível interação de Pressão de injeção de gás com a distancia da câmera. Um novo fator será considerado com o objetivo de reduzir ruídos na leitura da imagem, uma manta de sobrevivência para evitar o reflexo da luz IR de forma indevida.

### **Estratégia de experimentação**

Nesse experimento foram estudados 3 fatores em 2 níveis diferentes, conforme tabela 9: Fatorial completo de  $2^3$ , com 8 combinações possíveis (testes). Um fator novo foi introduzido, Manta, para verificar se reduz o reflexo do IR, melhorando a visualização da imagem da temperatura. Os fatores Pressão de Injeção do gás e Distância da câmera para o vazamento, tiveram seus níveis alterados, em relação ao DOE#1.

Alguns fatores foram considerados constantes durante o experimento (tipo de gás He, Tipo de câmera, Tamanho do vazamento 0,3mm<sup>2</sup>).

Tabela 9 - Fatores e níveis para o DOE#2

Fatores	Nível (-)	Nível (+)	Teoria
A – Manta	Sem	Com	O uso dessa manta reduz ruídos na reflexão de IR
B - Pressão de injeção do gás He	5 bars	9 bars	Esse fator influencia no contraste da temperatura
C - Distância da câmera	10 cm	45 cm	Esse fator influencia no contraste da temperatura
D – Tipo de gás	He		Durante esse experimento o gás utilizado será somente o He
E – Tipo de câmera	Modelo Optris 640		Durante esse experimento será utilizado apenas um modelo de câmera
F – Tamanho do vazamento	Furo pequeno, posição externa, na curva (0,3mm <sup>2</sup> )		Esse é o menor vazamento do DOE#1

Tabela 10 - Variáveis de fundo ou ruído e métodos de controle

Variáveis de ruídos	Métodos de controle
Fluxo de ar ambiente	Controlar pessoas se movimentando na área do experimento
Temperatura ambiente	Monitoramento
Setups dos níveis dos fatores	Monitoramento/Observar restrições
Emissividade do local (IR)	Utilizar uma manta de sobrevivência, que ajuda a minimizar a reflexão de infravermelhos do produto em teste. Neste experimento essa manta está sendo utilizada como um fator

O experimento foi conduzido em laboratório e está representado pela árvore de amostragem conforme figura 37.

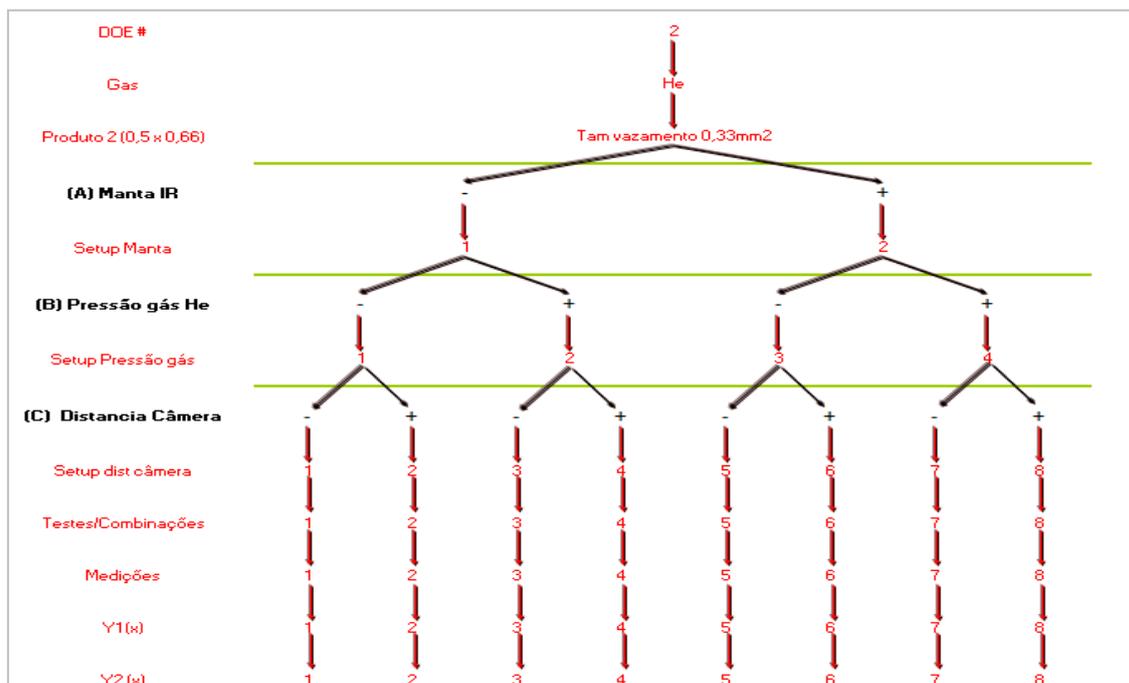


Figura 37 - Árvore de amostragem do experimento (DOE#2)

## Matriz de experimentação, coleta e análise dos dados - DOE#2

A matriz de experimentação, apresentada na tabela 11, mostra 8 combinações (testes) diferentes, variando os fatores A (Manta), B (Pressão de injeção de gás) e C (Distância da camera), conforme seus níveis (-1 e +1) que na prática correspondem aos níveis descritos na parte superior da tabela. As colunas A, B e C são a parte da matriz que direciona a manipulação dos níveis dos fatores durante o experimento e os valores da variável resposta, nas colunas Y1 e Y2 são observados e coletados.

Tabela 11 - Matriz de experimentação e coleta de dados (DOE#2)

Níveis									
(-1)	Sem	5 Bar	10cm						
(+1)	Com	9 Bar	45cm	Interações					
Testes	(A) Manta	(B) Pressão de injeção do gás	(C) Distância da câmara	AB	AC	BC	ABC	Y1 (Tempo p reduzir 0,4°C -seg)	Y2 (Tempo para visualizar o vazamento - seg)
1	-1	-1	-1	1	1	1	-1	20	14
2	-1	-1	1	1	-1	-1	1	20	13
3	-1	1	-1	-1	1	-1	1	8	12
4	-1	1	1	-1	-1	1	-1	5	13
5	1	-1	-1	-1	-1	1	1	13	17
6	1	-1	1	-1	1	-1	-1	20	18
7	1	1	-1	1	-1	-1	-1	18	17
8	1	1	1	1	1	1	1	17	15
Efeito (Y1)	3,8	-6,3	0,8	7,3	2,3	-2,8	-1,3		
Efeito (Y2)	3,8	-1,3	-0,3	-0,3	-0,3	-0,3	-1,3		

As colunas de interações são construídas a partir das colunas dos fatores principais (A, B e C) e portanto, não são manipuláveis. No entanto, com esse experimento é possível estudar sobre o efeito dos Fatores Principais e as interações possíveis entre eles (AB, AC, BC e ABC) e a magnitude dos efeitos pode ser calculada com base nos dados da variável resposta Y e observada na parte inferior da tabela.

### Medição da variável resposta (Y1 e Y2)

O método de medição de Y1 (tempo para reduzir a temperatura em 0,5°C na região do vazamento) e Y2 (tempo para visualizar a imagem termográfica do vazamento) são os mesmos adotados no DOE#1.

### Análise prática

Nesse experimento, foi utilizado apenas um produto com o furo de 0,3mm<sup>2</sup>, o menor do DOE#1, a pressão de injeção do gás foi aumentada (B) e adicionado uma manta para reduzir reflexo de IR (A).

Os resultados dos testes variaram entre 5s e 20s ou mais, conforme figura 38.

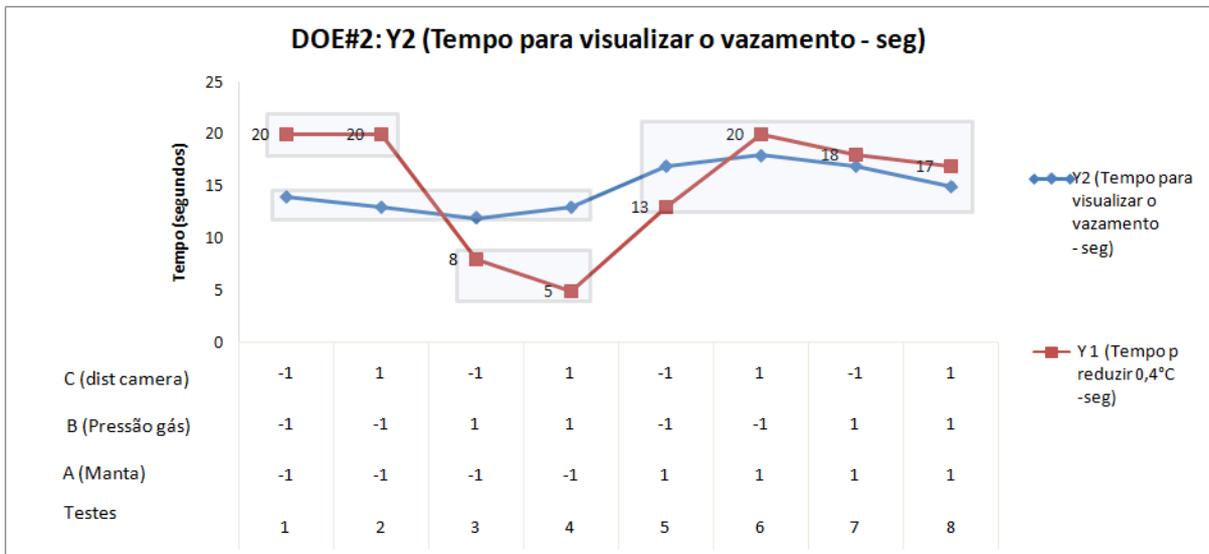


Figura 38 - Gráfico de tempo para detecção do vazamento (DOE#2)

As duas variáveis respostas, Y1 e Y2, discordaram entre si nos primeiros 4 testes, sem uso da Manta (A-), e praticamente concordam nos 4 últimos testes, com uso da Manta (A+).

Para Y1 e Y2, o uso da Manta (A+) aumentou o tempo de detecção.

Para o Y1, sem a manta (A-) e pressão do gás mais elevada (B+) o tempo de detecção reduz consideravelmente, ou seja, não foi observado um grande benefício em adicionar a manta.

### Análise gráfica

A análise gráfica é feita com auxílio gráfico, conforme figura 39, onde é apresentado a magnitude dos efeitos principais e interações que impactam na variável resposta Y1 e Y2, quando os fatores mudam seus níveis de -1 para +1.

O fator A (Manta) provocou um aumento no tempo de detecção para Y1 e Y2, com maior intensidade para o Y2, complementando o que foi visto na análise prática.

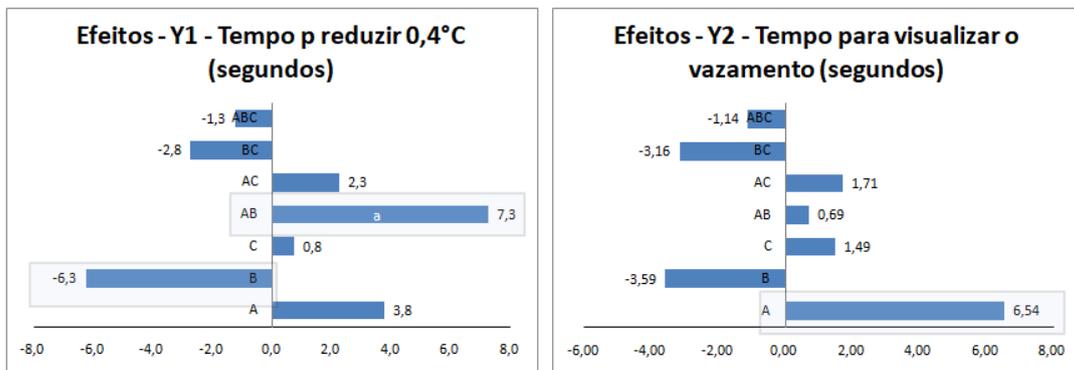


Figura 39- Gráfico de efeitos principais e interações (DOE#2)

O fator B (pressão do gás) provocou uma redução no tempo para detecção para Y1 e Y2, com maior intensidade para Y1.

A interação AB também mostra um aumento no tempo de detecção, que pode ser visto melhor no gráfico de interação AB na figura 39.

Nas figuras 40 e 41, os gráficos de efeitos principais, permitem uma melhor visualização do impacto de cada fator, quando variam do nível ( - ) para o nível ( + ).

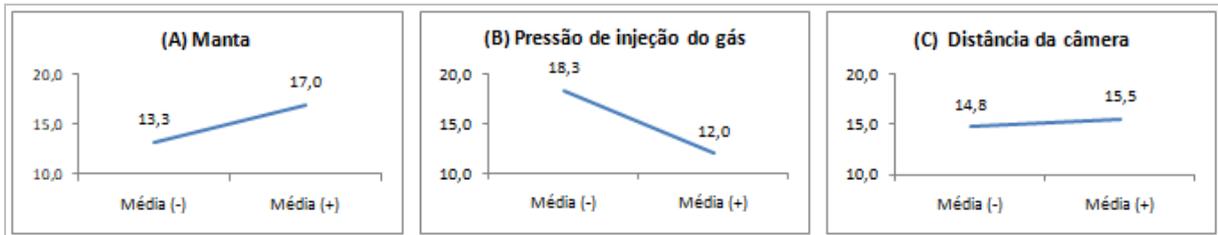


Figura 40 - Gráficos de efeito principal para Y1 (tempo para reduzir 0,5°C) – DOE#2

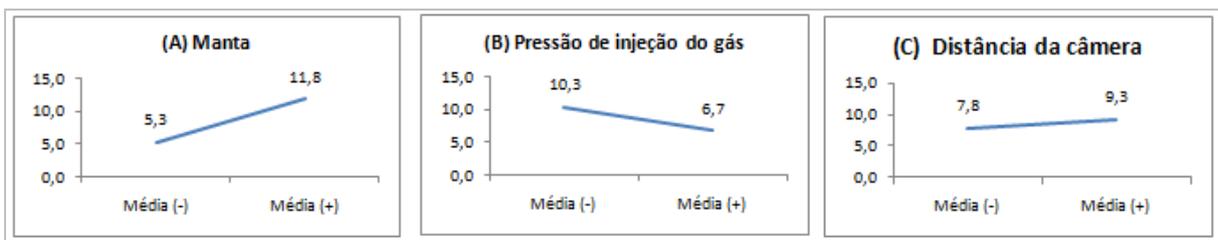


Figura 41 - Gráficos de efeito principal para Y2 (tempo para visualizar a imagem do vazamento)

Fator A: Uso da Manta (A+) leva mais tempo para ser detectado, nos dois Y's;

Fator B, o produto com pressão maior (B+) leva menos tempo para ser detectado para o Y1 e não significativo para o Y2, confirmando a mesma direção do DOE#1;

O fator C não apresentou variação significativa com a posição da camera próxima ou distante do ponto de vazamento para os dois Y's, se mantendo o mesmo resultado do DOE#1

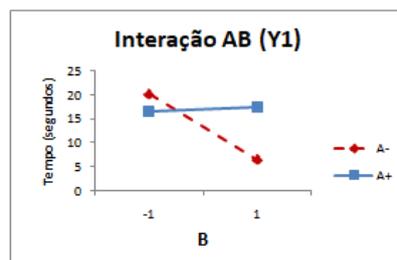


Figura 42 - Gráficos de interação AB para Y1 (tempo para reduzir 0,5°C) – DOE#2

A interação AB para o Y1, figura 42, mostra que o resultado depende do nível em que os dois fatores se encontram. Para se obter um resultado de tempo menor, o fator B deve estar no nível B+ (pressão 9Bar) e o fator A deve estar no nível A- (sem manta). Quando a pressão estiver em 5Bar (B -), usando a manta ou não (A - ou A+), o tempo para detecção é maior, se aproximando do tempo limite da linha de produção.

### **Aprendizados e próximos passos**

Esse experimento proporcionou os aprendizados abaixo:

- 1) O uso da manta para melhorar a detecção não produziu o que se esperava, pois provocou um aumento no tempo;
- 2) A pressão de injeção do gás He (B) maior continua provocando uma redução no tempo de detecção, aumentando a confiança nesse fator;
- 3) A distância da câmera para o ponto de vazamento (C) continua se mostrando pouco significativo.

### **4.3 PDSA #3**

#### **Objetivo:**

Avançar nos estudos para avaliar fatores de processo e método de medição com a camera de IR para identificar vazamento de gás em um produto de ar condicionado com possibilidades de aplicação em processo de fabricação em larga escala, dentro do tempo de 20 segundos.

#### **Informações e conhecimento prévio adquiridos sobre o assunto**

Até o momento, foi possível aprender sobre o comportamento dos fatores conforme abaixo:

- Tamanho do furo – Quanto menor o furo, mais tempo leva p detetar e mais difícil de visualizar com a câmera;
- Pressão de injeção He – Quanto maior a pressão, menos tempo leva para detetar o vazamento;
- Uso de manta para reduzir reflexo de IR – Não houve melhoria no tempo de detecção ao aplicar a manta enclausurando o produto;

O DOE#3 foi praticamente uma repetição do DOE#2, utilizando um produto com tamanho de vazamento menor, produzido artificialmente.

## Estratégia de experimentação

Neste experimento foram estudados 3 fatores em 2 níveis diferentes, conforme tabela 12: Fatorial completo de  $2^3$ , com 8 combinações possíveis (testes). Os fatores e níveis são os mesmos adotados no DOE#2. Dos fatores que permanecem constante durante o experimento foi adotado um tamanho de furo menor, reduzindo de 0,3mm<sup>2</sup> (DOE#2) e para 0,198mm<sup>2</sup> (DOE#3). O tipo de gás He e Tipo de câmera também se mantiveram constantes.

Tabela 12 - Fatores e níveis para o DOE#3

Fatores	Nível (-)	Nível (+)	Teoria/notas
A – Manta	Sem	Com	O uso dessa manta reduz ruídos na reflexão de IR
B - Pressão de injeção do gás He	5 bars	9 bars	Esse fator influencia no contraste da temperatura
C - Distância da câmera	10 cm	45 cm	Esse fator influencia no contraste da temperatura
D – Tipo de gás	He		Durante esse experimento o gás utilizado será somente o He
E – Tipo de câmera	Modelo Optris 640		Durante esse experimento será utilizado apenas um modelo de câmera
F – Tamanho do vazamento	Furo pequeno, posição externa, na curva, produzido artificialmente (0,198mm <sup>2</sup> )		Quanto menor o furo do vazamento, mais difícil de detectar

Tabela 13 - Variáveis de fundo ou ruído e métodos de controle

Variáveis de ruídos	Métodos de controle
Fluxo de ar ambiente	Controlar pessoas se movimentando na área do experimento
Temperatura ambiente	Monitoramento
Setups dos níveis dos fatores	Monitoramento/Observar restrições
Emissividade do local (IR)	Utilizar uma manta de sobrevivência, que ajuda a minimizar a reflexão de infravermelhos do produto em teste. Neste experimento essa manta está sendo utilizada como um fator

O experimento foi conduzido em laboratório e está representado pela árvore de amostragem (figura 43).

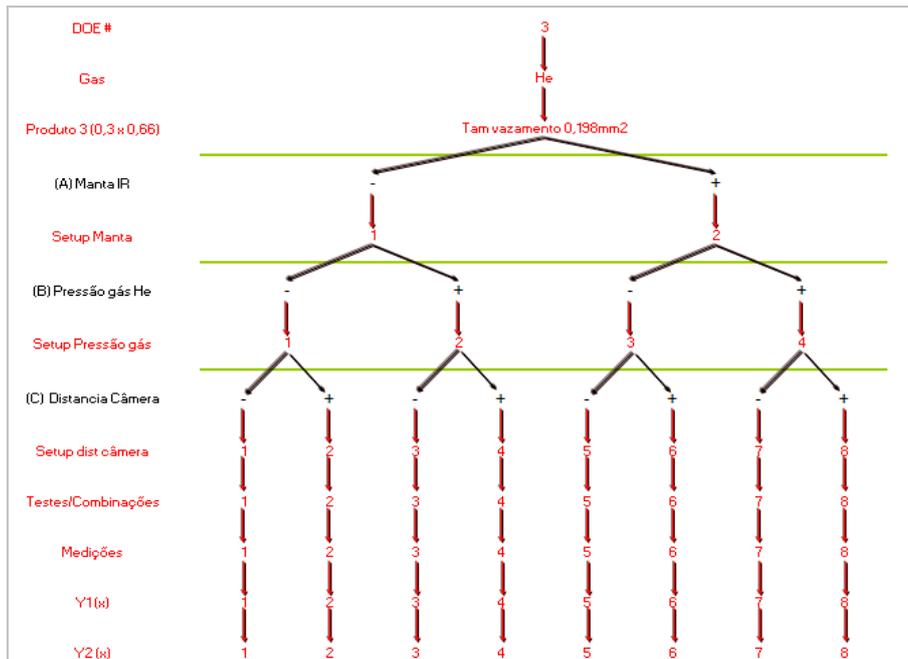


Figura 43 - Árvore de amostragem do experimento (DOE#3)

### Matriz de experimentação, coleta e análise dos dados - DOE#3

O experimento foi realizado, porém, com esse tamanho de furo do vazamento de 0,198mm<sup>2</sup>, produzido artificialmente, não foi possível identificar o vazamento através da câmera térmica, pois não produziu uma temperatura mais baixa na região do furo, o que parece ter ultrapassado o limite de detecção por imagem. Portanto, não houve coleta de dados e análise.

## 4.4 PDSA #4

### Objetivo:

Avançar nos estudos para avaliar fatores de processo e método de medição com a camera de IR para identificar vazamento de gás em um produto de ar condicionado com possibilidades de aplicação em processo de fabricação em larga escala, dentro do tempo de 20 segundos.

### Informações e conhecimento prévio adquiridos sobre o assunto

Até o momento, foi possível aprender sobre o comportamento dos fatores conforme abaixo:

- Tamanho do furo – Quanto menor o furo, mais tempo leva p detectar e mais difícil de visualizar com a câmera; No DOE#3, o teste com um furo de 0,198mm<sup>2</sup> não foi possível detectar o vazamento com a câmera.

- Pressão de injeção He – Quanto maior a pressão, menos tempo leva para detetar o vazamento;
- Uso de manta para reduzir reflexo de IR – Não houve melhoria no tempo de detecção ao aplicar a manta enclausurando o produto;

O DOE#4 será praticamente uma repetição do DOE#2, utilizando produtos com tamanho de vazamentos menores e será trocado o tipo de gás de He para Refrigerante R410 (gás utilizado dentro do sistema de refrigeração do ar condicionado)

### Estratégia de experimentação

Neste experimento foram estudados 3 fatores em 2 níveis diferentes, conforme tabela 14: Fatorial completo de  $2^3$ , com 8 combinações possíveis (testes). Considerando que o limite de detecção com a camera foi atingido no DOE#3, para furo de 0,198mm<sup>2</sup> e utilizando o gasHe, no DOE#4 está sendo planejado usar o gás refrigerante R410 e furos de 0,198mm<sup>2</sup> ou menor ainda.

Tabela 14 - Fatores e níveis para o DOE#4

Fatores	Nível (-)	Nível (+)	Teoria
A – Tamanho do vazamento	0,198mm <sup>2</sup>	0,105mm <sup>2</sup>	Quanto menor o furo do vazamento, mais difícil de detectar
B – Manta	Sem	Com	O uso dessa manta reduz ruídos na reflexão de IR
C - Distância da câmera	10 cm	45 cm	Esse fator influencia no contraste da temperatura
D – Tipo de gás	R410		Durante esse experimento o gás utilizado será somente o R410
E – Tipo de câmera	Modelo Optris 640		Durante esse experimento será utilizado apenas um modelo de câmera

Tabela 15 - Variáveis de fundo ou ruído e métodos de controle

Variáveis de ruídos	Métodos de controle
Fluxo de ar ambiente	Controlar pessoas se movimentando na área do experimento
Temperatura ambiente	Monitoramento
Setups dos níveis dos fatores	Monitoramento/Observar restrições
Emissividade do local (IR)	Utilizar uma manta de sobrevivência, que ajuda a minimizar a reflexão de infravermelhos do produto em teste. Neste experimento essa manta está sendo utilizada como um fator

O experimento foi conduzido em laboratório e está representado pela árvore de amostragem (figura 41).

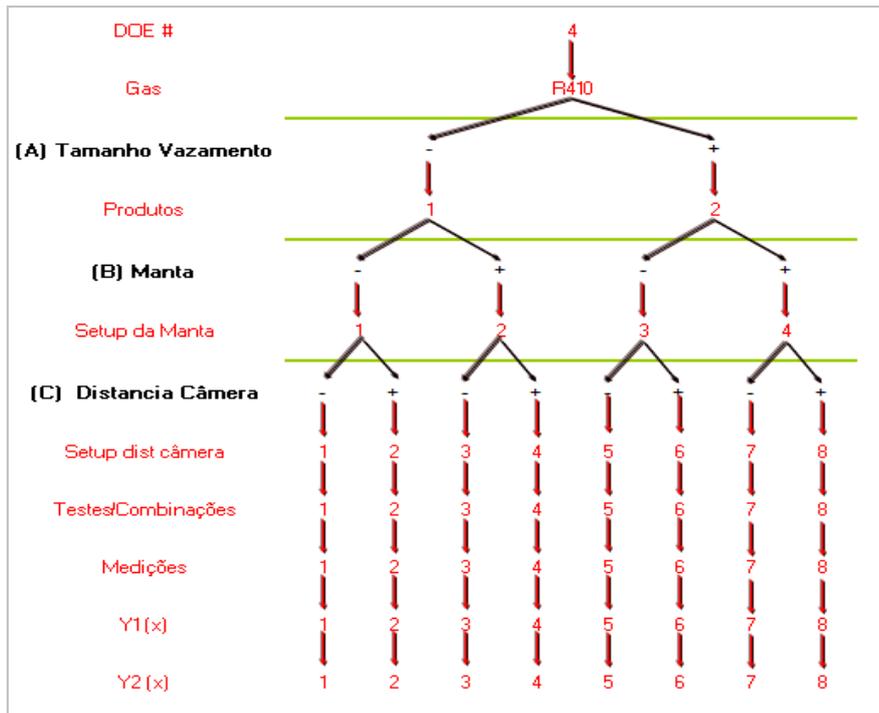


Figura 44 - Árvore de amostragem do experimento (DOE#4)

### Matriz de experimentação, coleta e análise dos dados - DOE#4

A tabela 16 apresenta a matriz de experimentação, com 8 combinações (testes) diferentes, variando os fatores A (Tamanho do vazamento), B (Manta) e C (Distância da camera), conforme seus níveis (-1 e +1) que na prática correspondem aos níveis descritos na parte superior da tabela. As colunas A, B e C são a parte da matriz que direciona a manipulação dos níveis dos fatores durante o experimento e os valores da variável resposta, nas colunas Y1 e Y2 são observados e coletados.

Tabela 16 - Matriz de experimentação e coleta de dados (DOE#4)

Níveis									
(-1)	(0,66 x 0,33 mm)	Sem	10cm						
(+1)	(0,7 x 0,15 mm)	Com	45cm	Interações					
Testes	(A) Tamanho do vazamento	(B) Manta	(C) Distância câmera	AB	AC	BC	ABC	Y1 (Tempo p reduzir 0,5°C -seg)	Y2 (Tempo para visualizar o vazamento -seg)
1	-1	-1	-1	1	1	1	-1	4	6
2	-1	-1	1	1	-1	-1	1	7	4
3	-1	1	-1	-1	1	-1	1	8,3	8
4	-1	1	1	-1	-1	1	-1	9,1	9
5	1	-1	-1	-1	-1	1	1	3,5	2
6	1	-1	1	-1	1	-1	-1	3,5	5
7	1	1	-1	1	-1	-1	-1	5	5
8	1	1	1	1	1	1	1	3,5	5
	0	0	0	0	0	0	0		
Efeito (Y1)	-3,2	2,0	0,6		-1,2	-1,3	-0,9		
Efeito (Y2)	-2,5	2,5	0,5		-1,0	1,0	0,0		

## Medição da variável resposta (Y1 e Y2)

O método de medição de Y1 (tempo para reduzir a temperatura em 0,5°C na região do vazamento) e Y2 (tempo para visualizar a imagem termográfica do vazamento) são os mesmos adotados no DOE#1.

## Análise prática

Foi possível detectar o vazamento de tamanho de furos pequenos, não identificados com o uso do gás He no DOE#3. O uso de gás refrigerante R410 provocou uma diferença de temperatura mais significativa, possibilitando a visualização do vazamento por imagem termográfica com temperatura mais baixa na região do vazamento;

Os resultados dos testes variaram entre 2s e 9s, melhorando muito o tempo para detecção, com uma folga razoável em relação ao tempo Max permitido de 20s, conforme figura 45.

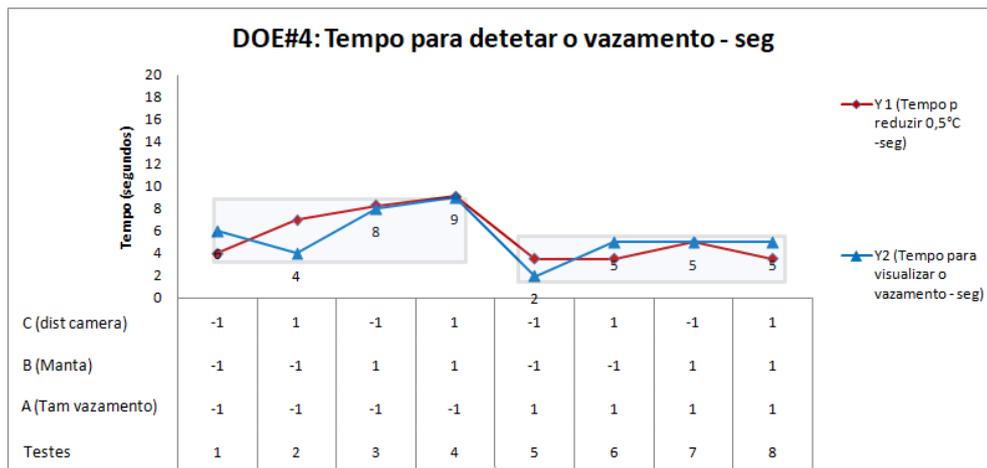


Figura 45 – Gráfico de tempo para detecção do vazamento (DOE#4)

Os dois furos são pequenos, na ordem de 3g/ano e com formatos diferentes, o que foi observado formas de visualização diferentes. O furo com formato redondo (A-), foi possível identificar o gás vazando e o furo com formato retangular (A+), a área do furo ficou mais fria, destacando bem a temperatura em relação as demais áreas;

As duas variáveis respostas, Y1 e Y2, praticamente concordam entre si em todos os testes;

O fator A (tamanho do vazamento) é o que mostra maior variação quando muda de (A-) para (A+). Seu formato pode influenciar no tempo de detecção, porém ambos levaram tempos bem pequenos para serem detetados;

O uso da Manta (B+) não mostrou-se significativa para redução do tempo

## Análise gráfica

A figura 46 mostra a magnitude dos efeitos principais e interações que impactam na variável resposta Y1 e Y2, quando os fatores mudam seus níveis de -1 para +1.

O fator A (Tamanho do vazamento) foi o que provocou maior variação quando muda de -1 para +1, uma redução no tempo de detecção, complementando o que foi visto na análise prática;

O fator B (Manta) foi a segunda maior variação quando muda de -1 para +1, provocando um aumento no tempo de detecção quando a manta é usada (+1);

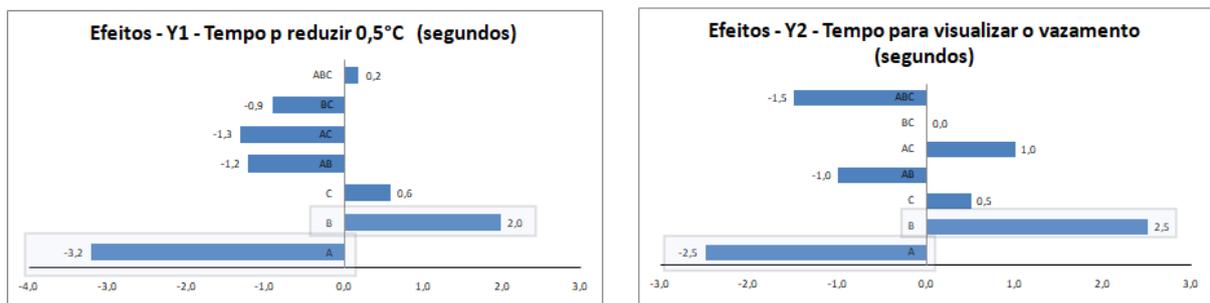


Figura 46 – Gráfico de efeitos principais e interações (DOE#4)

Observando a figura 47 e 48, verifica-se o mesmo comportamento em Y1 e Y2:

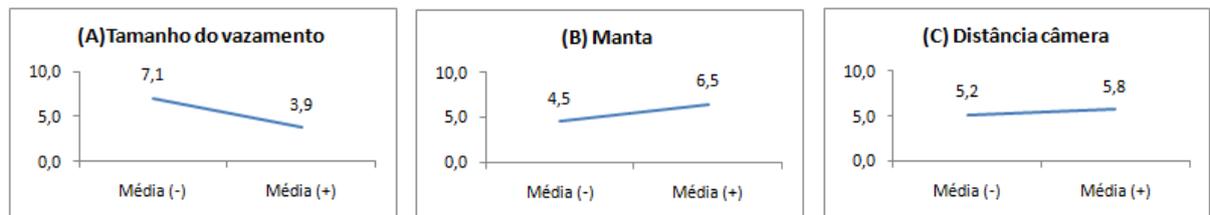


Figura 47 – Gráficos de efeito principal para Y1 (tempo para reduzir 0,5°C) – DOE#4

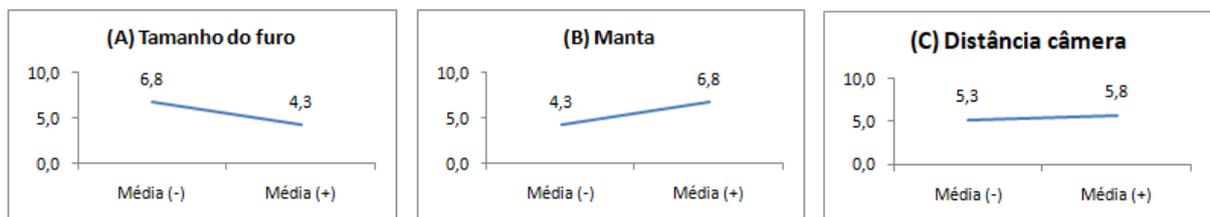


Figura 48 – Gráficos de efeito principal para Y2 (tempo para visualizar a imagem do vazamento) – DOE#4

Fator A: Tamanho do vazamento (A-) leva mais tempo para ser detetado

Fator B: Uso da manta (B+) aumenta o tempo de detecção, confirmando o mesmo aprendizado dos experimentos anteriores, aumentando a confiança de que a manta não traz benefícios significativos.

Fator C não apresentou variação significativa com a posição da camera próxima ou distante do ponto de vazamento, se mantendo o mesmo resultado dos experimentos anteriores.

### **Aprendizados e próximos passos**

Esse experimento trouxe os aprendizados abaixo:

- 1) O uso do gás refrigerante R410 possibilitou detectar o vazamentos muito pequenos (0,198mm<sup>2</sup> e 0,115mm<sup>2</sup> – na ordem de 3g/ano) através de imagem termográfica, por ter produzido uma diferença de temperatura maior na região do vazamento. O que não foi possível identificar no DOE#3 quando utilizado o gás He
- 2) O uso da manta não traz benefícios de melhoria na leitura da imagem;
- 3) A distância da câmera para o ponto de vazamento (C) se mostra pouco significativa.

## **5 Aprendizagem de máquina para identificação de vazamento**

Nesse capítulo é descrito o método de coleta, preparação e utilização de dados adotando técnicas de aprendizagem de máquina supervisionado com algoritmos que permitem a avaliação quanto sua performance em classificar regiões de produtos com vazamento e sem vazamento, utilizando o software Orange.

### **5.1 Estratégia de amostragem (recolha de dados)**

Para o estudo de aprendizagem de máquina foram utilizados os dados coletados no último experimento descrito no capítulo anterior (DOE#4), conforme árvore de amostragem, na figura 49, onde foram realizados 8 testes com suas respectivas combinações dos fatores A, B e C. Tomando como exemplo, o teste 1 foi realizado considerando o gás R410, com tamanho de furo 0,22mm<sup>2</sup>, utilizando o produto A, sem manta de infravermelho e a distância da câmera perto (10cm) do produto em teste. O teste 2 foi similar ao teste 1, mudando apenas a distância da câmera para longe (45cm) em relação ao produto em teste. Dessa forma, 8 testes foram realizados e coletados os dados de temperatura da superfície do trocador de calor.

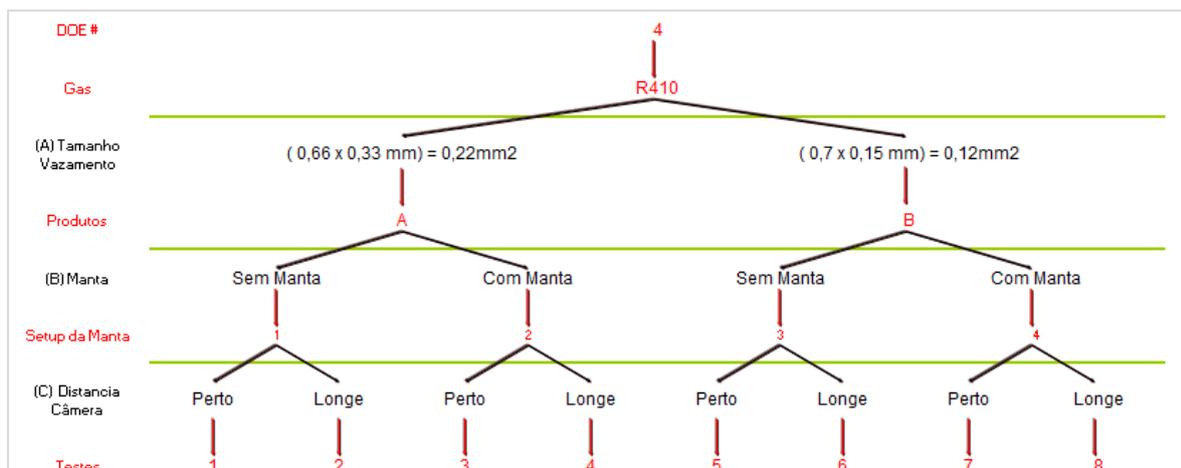


Figura 49 – Árvore de amostragem – DOE#4

Para cada teste, um vídeo foi gravado com a câmera térmica, registrando a temperatura do objeto presente na gravação, por um período de 20 segundos. Um software que acompanha a câmera, possibilita visualizar o vídeo e exportar para uma planilha em Excel os dados de temperatura selecionada na imagem. Foram selecionados 8 pontos (P1 a P8) para coleta de temperatura, sendo 1 deles o ponto onde se encontra o vazamento, conforme figura 50.

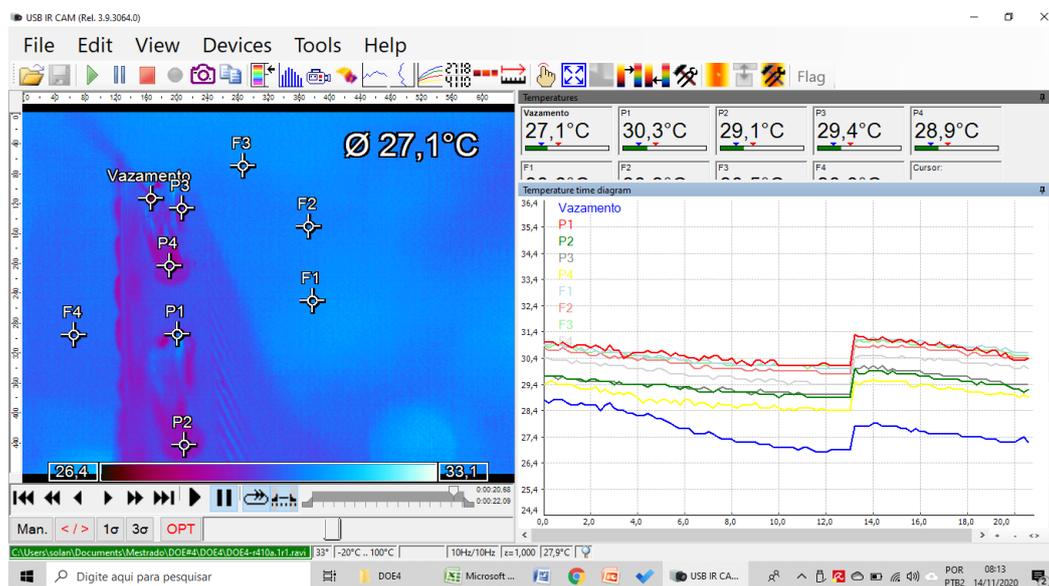


Figura 50 – Pontos de coleta de temperatura através da imagem de IR

Para cada ponto (P1 a P8), foram coletados os dados de temperatura a cada 0,2s por um período de 20s, resultando na estratégia da figura 51.

Dos 8 testes, 4 foram utilizados para treino (destacado em azul) e 4 foram utilizados para teste, selecionados de forma aleatória.

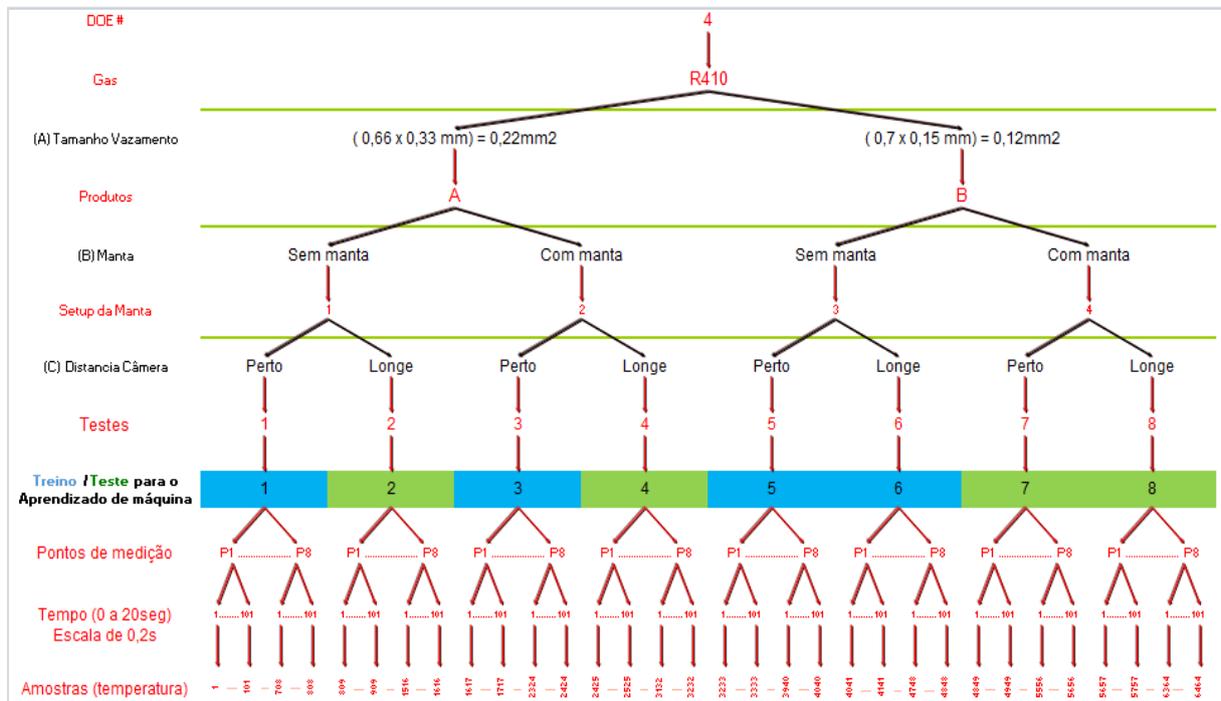


Figura 51 – Estratégia de amostragem de TREINO e TESTE / Amostras de temperatura para cada experimento

Os dados coletados foram agrupados em duas planilhas dentro de um arquivo em Excel: Uma planilha com os dados para TREINO e uma planilha com dados para TESTE. Na tabela de treino, uma coluna foi adicionada para classificar a amostra como “Vazamento” ou “Normal”, que foi usada para identificar a área/ponto com vazamento e sem vazamento durante o processo de aprendizagem supervisionado. A tabela 17 apresenta uma amostra da tabela de treino, com os atributos nas colunas e as instâncias nas linhas. Os atributos considerados foram os fatores o Produto (A e B), fator A (tamanho do vazamento), fator B (uso de manta IR), fator C (distância da câmera para o produto) e a Temperatura ao longo do tempo (0s a 20s, medidos a cada 0,2s), resultando em 106 atributos/features. A coluna “Classificação” foi o atributo considerado para treino do algoritmo, indicando o ponto com “Vazamento” e “Normal”.

Tabela 17 - Dados de Treino (amostra dos dados)

<i>Atributos</i>	<i>Produto</i>	<i>(A) Tamanho do vazamento</i>	<i>(B) Manta IR</i>	<i>(C) Distância da câmera</i>	<i>Classificação</i>	<i>Posição da Medição</i>	<i>Temp. (t 0)</i>	<i>Temp. (t 0,2)</i>	<i>Temp. (t 0,4)</i>
INSTÂNCIAS/ EXEMPLOS	B	0,12mm2	sem	10cm	Vazamento	P1	27,4	27,4	27,3
	B	0,12mm2	sem	10cm	Normal	P2	27,9	28	27,8
	B	0,12mm2	sem	10cm	Normal	P3	28,2	28,2	28,2
	B	0,12mm2	sem	10cm	Normal	P4	28,1	28,1	27,9
	B	0,12mm2	sem	10cm	Normal	P5	27,9	27,9	27,8
	B	0,12mm2	sem	10cm	Normal	P6	30	30	30,1
	B	0,12mm2	sem	10cm	Normal	P7	27,4	27,3	27,2
	B	0,12mm2	sem	10cm	Normal	P8	28,9	28,9	28,9
	A	0,22mm2	com	10cm	Vazamento	P1	28,8	28,9	28,7
	A	0,22mm2	com	10cm	Normal	P2	28,5	28,5	28,5

O conjunto total de dados para treino e teste dos algoritmos encontram-se resumidos na tabela 18, com um total de 32 instâncias para treino e 32 instâncias para teste. Pode-se notar que os dados possuem um desbalanceamento entre as instâncias classificadas como “Normal” e “Vazamento”. Isso se deu a quantidade de amostras limitadas para o estudo.

18 – Tabela com total de dados para treino e teste dos algoritmos

	Dados de treino		Dados de teste		
	Normal	Vazamento	Normal	Vazamento	Total
Instâncias	28	4	28	4	64

## 5.2 Aprendizagem de máquina utilizando software Orange

Na figura 52 é apresentado o ambiente do Software Orange, onde é possível fazer um projeto de aprendizagem de máquina, utilizando os “widgets” ou “ferramentas” prontas para serem processadas. O projeto foi construído em 3 blocos: Bloco 1, carregamento de dados; Bloco 2, algoritmos de aprendizagem; Bloco 3, avaliação de performance dos algoritmos com base nos dados de entrada (teste e treino).

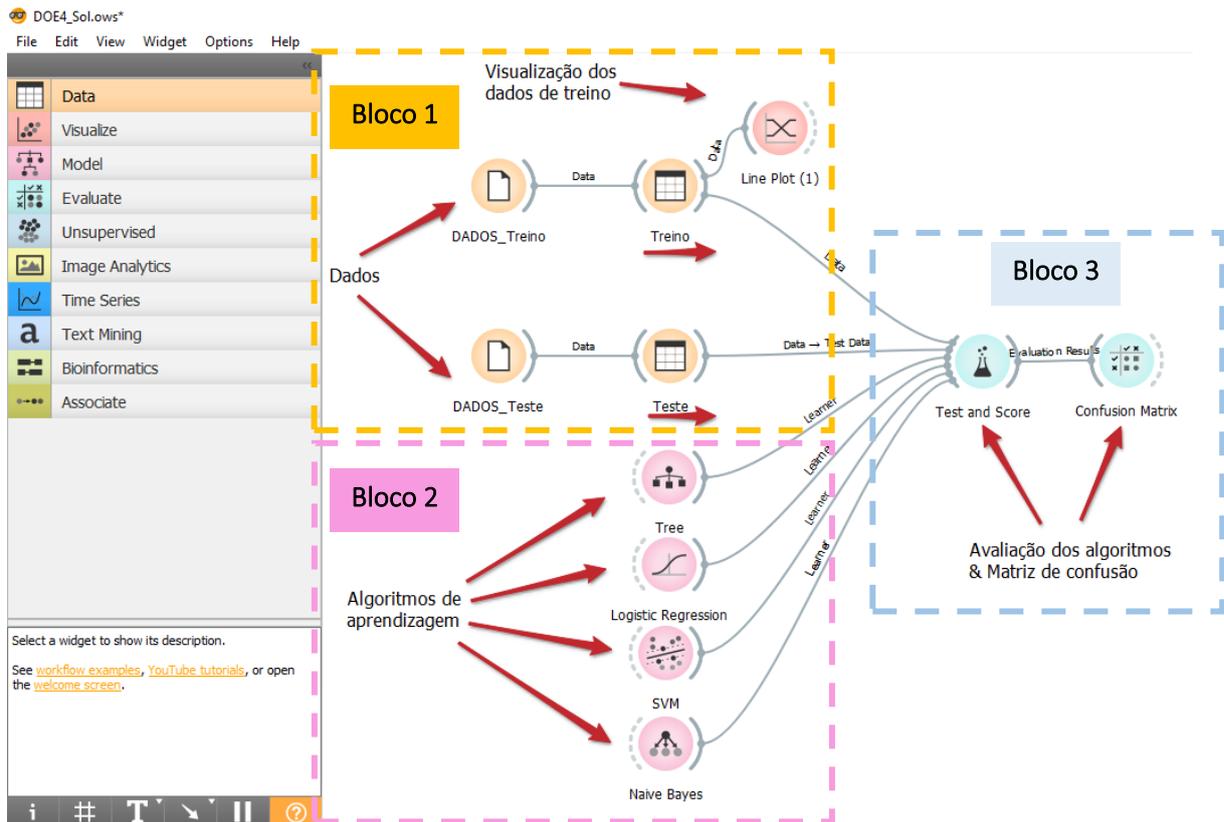


Figura 52 – Projeto de aprendizado de máquina para classificação de vazamento

### Bloco 1 – Carregamento de dados

Para carregar os dados de um arquivo, utilizou-se a ferramenta “file” arrastando para a área do “Canvas”. Clica duas vezes na ferramenta e uma janela se abre para fornecer a origem dos dados, conforme figura 53. Os dados são carregados automaticamente da planilha selecionada. O tipo (categóricas, numéricas, texto) e a função (feature, target) das variáveis devem ser revisados e ajustados. Importante identificar a variável para treinamento de classificação de dado. Nesse caso, a linha “Classificação” com os valores “Vazamento” e “Normal” são ajustados como “target” na coluna “Role”, indicando o aprendizado supervisionado.

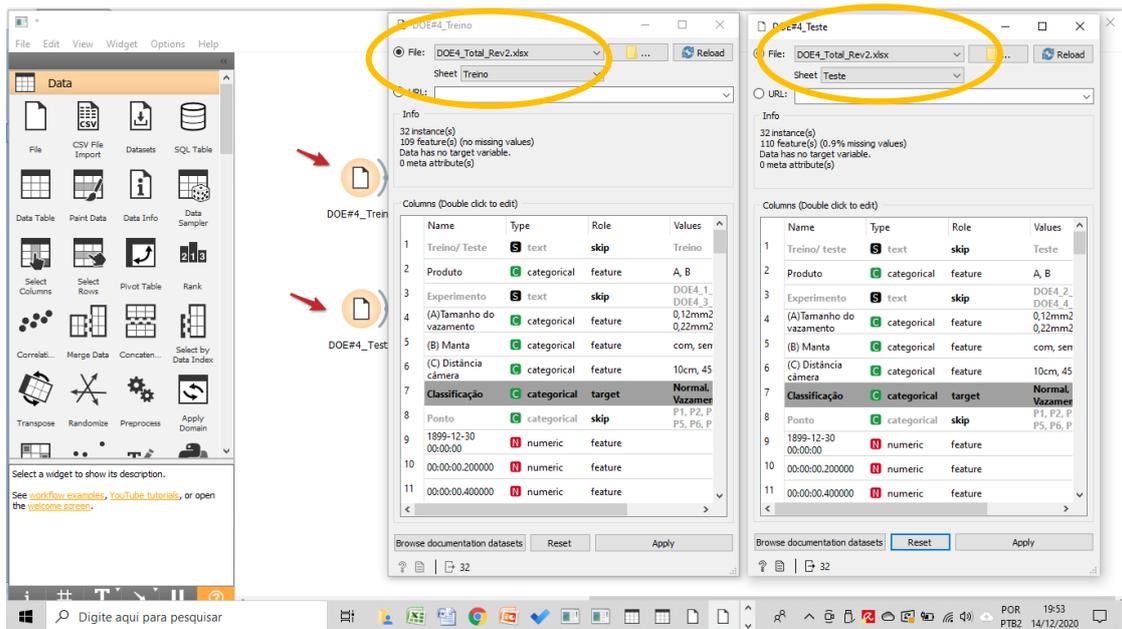


Figura 53 – Carregamento de dados de um arquivo Excel e ajustes dos tipos de dados e funções

O carregamento de dados (figura 53) está relacionado com os dados de “Treino”, identificados na planilha “Treino” e o carregamento para os dados de “Teste”, identificado no planilha “Teste”.

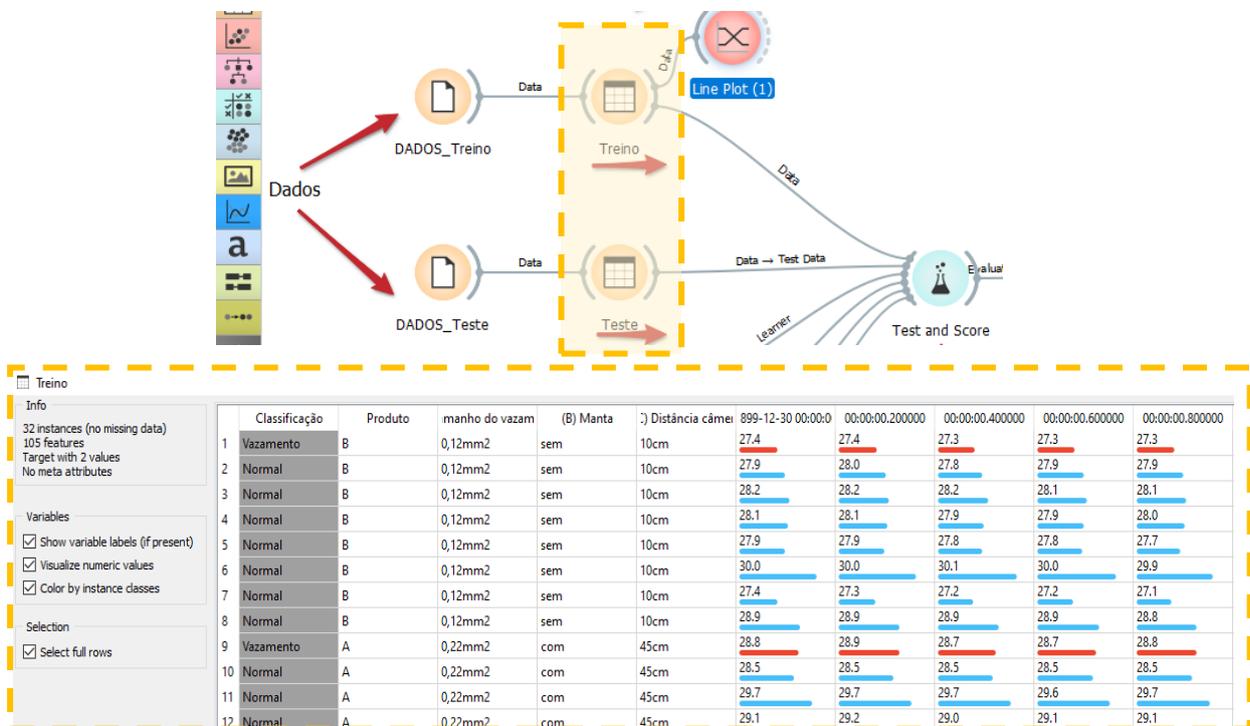


Figura 54 – Visualização da tabela com dados carregados

Arrastando o ícone da ferramenta “Data table” para o Canvas, renomeado para “Treino” e conectando a “Dados\_Treino”, é possível visualizar uma tabela com os dados carregados, conforme figura 54. Desta forma, os dados estão prontos para serem utilizados na ferramenta “Test & Score” que realiza a avaliação dos algoritmos.

É possível ainda visualizar graficamente os dados da tabela com a ferramenta “Line Plot” na figura 55, mostrando o comportamento da temperatura ao longo do tempo para os pontos classificados como “Normal” em azul e “Vazamento” em vermelho, que apresentam comportamento distintos em média, porém, com zonas de sobreposição.

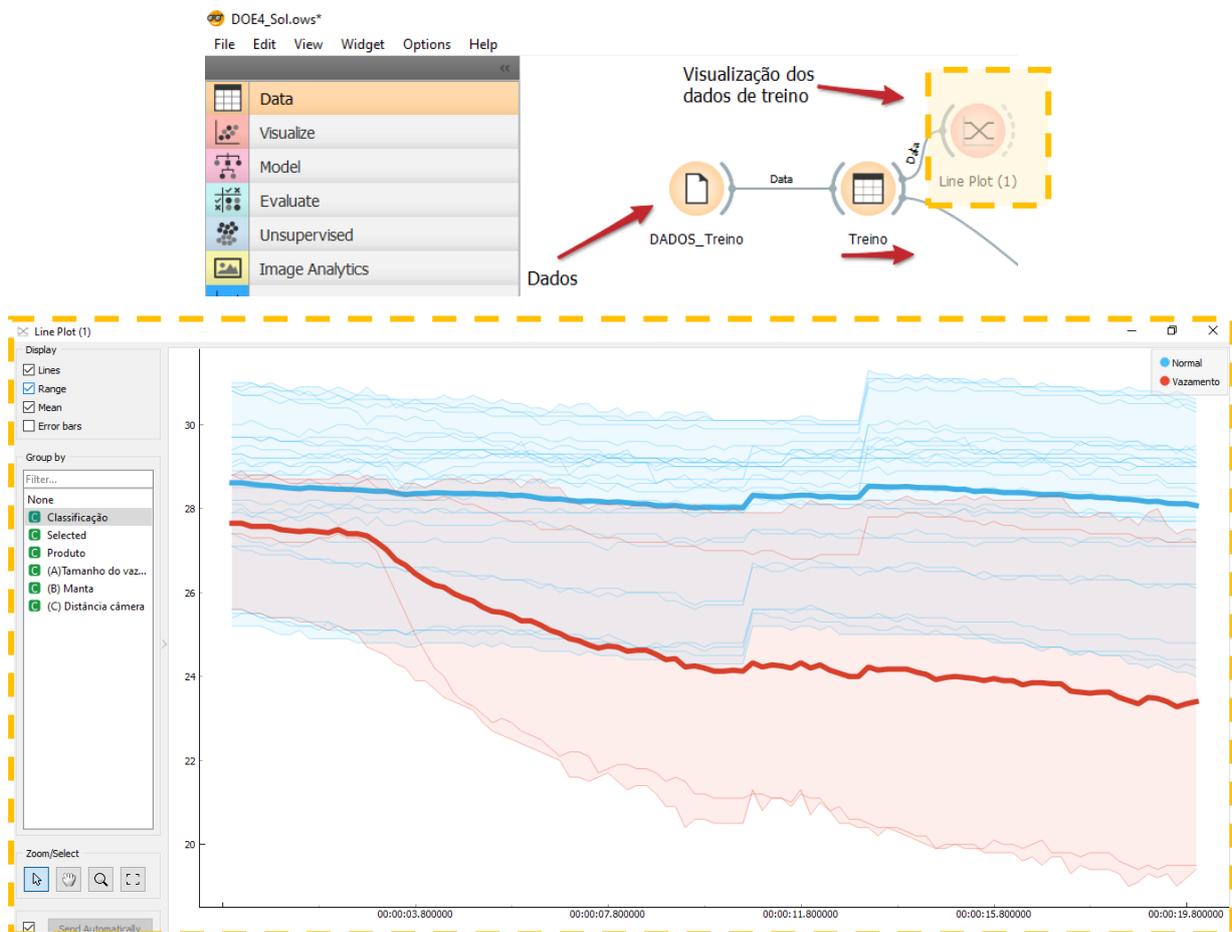


Figura 55 – Visualização grafica dos dados utilizando a ferramenta “Line Plot”

## Bloco 2 e 3 – Algoritmos de aprendizagem de máquina e avaliação de performance dos algoritmos

Os blocos 2 e 3 estão intrinsecamente ligados, uma vez que os dados estão prontos para o aprendizado, sendo necessário testar os algoritmos que mais tem aderência aos dados. Foram utilizados aqui 4 algoritmos que são utilizados para aprendizado supervisionado (Árvore de decisão, Regressão Logística, SVM e NaiveBayes), conforme figura 56.

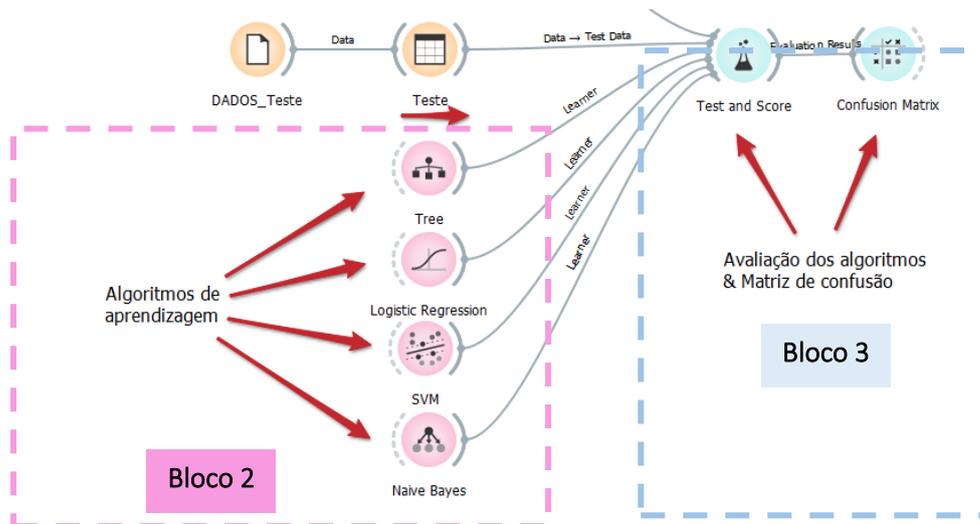


Figura 56 – Algoritmos de aprendizagem conectados ao teste de performance “Testand Score”

Para fazer a avaliação dos algoritmos, utilizou-se a ferramenta “Testand Score”, com dados de entrada (tabelas de Treino e Teste) e os algoritmos para treino avaliação de performance.

Os quatro algoritmos foram avaliados pela matriz de confusão e indicadores de performance que medem a habilidade de prever a classificação de uma região com “Vazamento” (VP – Verdadeiro Positivo) ou “Normal” (VN – Verdadeiro Negativo).

## 6 Resultados e discussões

Nesta seção serão apresentados os resultados dos experimentos sequenciais, conduzidos inicialmente para avaliar a possibilidade de detecção de vazamento de gás com uma câmera térmica em ambiente industrial. Na sequência, os resultados da aplicação de aprendizagem de máquina supervisionado utilizando os dados de temperatura capturados pela câmera de regiões com e sem vazamento em trocadores de calor de ar condicionados.

### Experimentos sequenciais

A figura 57 apresenta o resumo de 24 testes executados em 3 ciclos de experimentação (DOE#1, #2 e #4), com ajustes em diferentes condições, por exemplo, *tamanho do furo do vazamento, pressão do gás, distância da câmera, manta térmica, tipo de gás*. O último ciclo de testes apresentou menor tempo para detectar o vazamento. As conclusões principais após a análises destes experimentos são as seguintes:

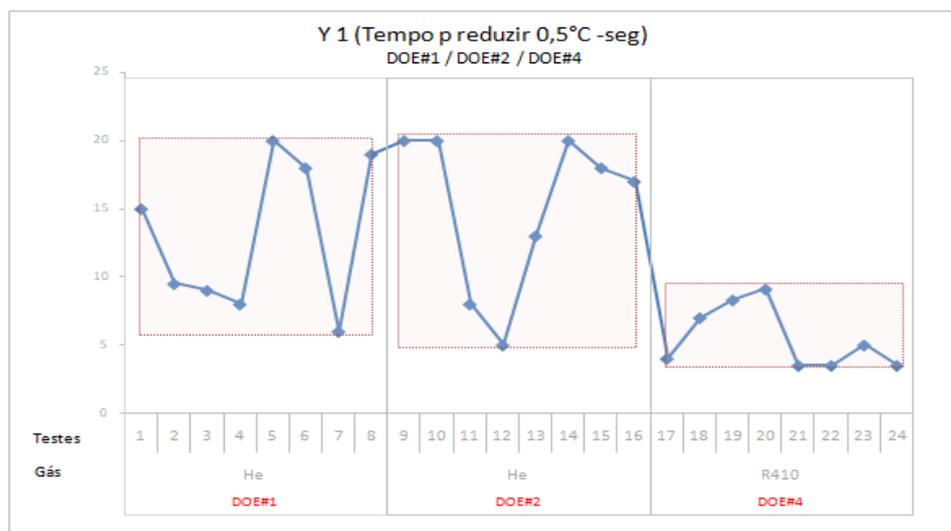


Figura 57– Resumo dos PDSAs para Y (tempo para detectar vazamento)

- O fator mais significativo para reduzir o tempo foi o tipo de gás, onde o R410 proporciona uma resposta mais rápida em relação ao He, atendendo o limite de processo de 20s;
- O primeiro e segundo ciclo de experimentos apresentaram variações muito grandes no tempo para detectar o vazamento, chegando até passar do limite de processo;
- A região de vazamento apresenta uma queda de temperatura ao longo do tempo e esse dado pôde ser utilizado para aprendizagem de máquina.

## Performance dos algoritmos

Este estudo inclui experiências com múltiplos algoritmos de machine learning, sobre os mesmos dados. Estes algoritmos tiveram desempenhos diferentes. É pertinente também analisar e discutir os desempenhos destes algoritmos. A tabela 19 apresenta uma comparação dos resultados de performance dos algoritmos treinados, considerando os indicadores de avaliação.

Tabela 19 - Avaliação dos algoritmos com dados de Treino e Teste

Algoritmo	Quanto maior o índice, melhor				Quanto menor o índice, melhor	
	CA – Classification Accuracy	Recall	Precision	F1	% FP	% FN
Árvore de decisão	0,84	0,00	0,00	0,00	0,03	0,13
SVM	0,91	0,25	1,00	0,40	0,00	0,09
Regressão logística	0,94	0,50	1,00	0,67	0,00	0,06
Naive Bayes	0,44	0,50	0,11	0,18	0,50	0,06
Leitura da avaliação	% de Acertos na classificação de VP e VN	% de Acertos de VP entre todos os VP	% de Acertos na classificação de VP	% na Ponderação entre Recall e Precisão	% de Falso Positivo	% de Falso Negativo
Fórmulas	$\frac{(VP+VN)}{(VP+VN+FP+FN)}$	$\frac{(VP)}{(VP+FN)}$	$\frac{(VP)}{(VP+FP)}$	$\frac{(2 \times \text{Precisão} \times \text{Recall})}{(\text{Precisão} + \text{Recall})}$	$\frac{(FP)}{(VP+VN+FP+FN)}$	$\frac{(FN)}{(VP+VN+FP+FN)}$

- Dos 4 algoritmos, a Árvore de Decisão obteve a pior resultado nos índices de Precisão, Recall e F1 com baixa performance. Seguido do Naive Bayes com um alto índice de FP, indicando que o produto está “Normal” e, no entanto, tem “Vazamento”
- Regressão Logística apresentou melhores resultados: 100% de Precisão na classificação de VP, zero classificação de FP. Tem oportunidade de melhoria nos indicadores de Recall, com 50% de acertos de defeitos reais (VP), afetado pela classificação de FN.
- Para aumentar a confiança nas previsões, seriam necessários mais dados de processo de ar-condicionado, principalmente de “Vazamento”, para que a avaliação utilize dados de VP e VN balanceados.

Os indicadores são extraídos da matriz de confusão, tabela 20, composta pela previsão de classificação do algoritmo comparados aos dados de teste (Real), com 4 tipos de previsões:

- Verdadeiro Positivo: Algoritmo diz que *tem defeito* “Vazamento” e *tem defeito*;
- Verdadeiro Negativo: Algoritmo diz que *não tem defeito* “Normal” e *não tem defeito*;

- Falso Positivo: Algoritmo diz que *tem defeito* “Vazamento” e *não tem defeito*;
- Falso Negativo: Algoritmo diz que *não tem defeito* “Normal” e *tem defeito*;

Tabela 20 - Matriz de confusão dos algoritmos

Algoritmo	Falso Positivo (FP)	Verdadeiro Positivo (VP) “VAZAMENTO”	Falso Negativo (FN)	Verdadeiro Negativo (VN) “NORMAL”	Total
Árvore de decisão	1	0	4	27	32
SVM	0	1	3	28	32
Regressão logística	0	2	2	28	32
Naive Bayes	16	2	2	12	32
Real		4		28	32

Os dados de teste são compostos por 28 instâncias de áreas sem vazamento “Normal” (VN) e 4 instâncias de áreas com “Vazamento” (VP). Com essa análise complementar utilizando a matriz de confusão, é possível observar que o algoritmo Árvore de decisão classificou os 4 “Vazamentos” como Falso Negativo (FN), assim como o Naive Bayes classificou 16 Falso Positivos, onde ambos tiveram uma baixa performance na classificação. Regressão logística classificou corretamente todos os Verdadeiros Negativos (VN) e acertou 50% dos Verdadeiros Positivos (VP).

## 7 Conclusões e trabalhos futuros

Com o decorrer do desenvolvimento do trabalho, foi possível observar a evolução do conhecimento sobre a utilização de uma câmera termográfica IR para identificar vazamento de gás em um sistema de ar condicionado, dentro do tempo de 20 segundos, e utilizar os dados de temperatura da superfície do produto, capturada pela câmera, para aprendizagem de máquina supervisionado, classificando uma região com “vazamento” ou região “normal”. Foi utilizado um software de código aberto, Orange, que permitiu a avaliação de algoritmos, onde o algoritmo Regressão Logística obteve melhor performance de classificação, com bom índice na classificação de Verdadeiros Negativos (Normal) e oportunidades de aprimoramento em trabalhos futuros. Para evolução na performance do algoritmo, é necessário a coleta de mais dados utilizando a câmera em amostras do processo produtivo de ar condicionados com Verdadeiros Positivos (regiões com vazamento). Contudo, este trabalho demonstrou que é possível fazer a detecção de vazamento de gás, de forma automática, utilizando algoritmos de aprendizagem, em ambiente industrial real, que é o contributo principal deste trabalho de

dissertação. Essa etapa do estudo é fundamental como subsídios para validação de um novo conceito de teste de vazamento de gás, dando uma segurança para a empresa seguir em frente com investimentos na fase desenvolvimento do software, hardware e integração da câmera ao processo prdutivo, com a finalidade de transferir o julgamento de um produto “Com Vazamento” e “Normal” para um sistema automático de detecção.

## 8 Referências Bibliográficas

- Alpaydin, E. (2010). *Introduction to Machine Learning* (M. I. of Technology (ed.); 2nd ed.). The MIT Press. <http://www.ncbi.nlm.nih.gov/pubmed/24272434>
- Camilo, C., & Silva, J. (2009). *Mineração de Dados: Conceitos, tarefas, métodos e ferramentas. Universidade Federal de Goiás (UFG)*, 29.
- Côrtes, S. D. C., Porcaro, R. M., & Lifschitz, S. (2002). *Mineração de Dados – Funcionalidades, Técnicas e Abordagens. PUC-Rio Informática*, 35. [ftp://ftp.inf.puc-rio.br/pub/docs/techreports/02\\_10\\_cortes.pdf](ftp://ftp.inf.puc-rio.br/pub/docs/techreports/02_10_cortes.pdf)
- Coutinho, C. P., Souza, A., Dias, A., Bessa, F., Ferreira, M. J., & Vieira, S. (2009). *Investigação\_Acção\_Metodologias.pdf* (p. 26).
- Coutinho, M. (2016). *Funcionamento básico de um ciclo de refrigeração*. <https://pt.linkedin.com/pulse/funcionamento-básico-de-um-ciclo-refrigeração-marcos-coutinho>
- Creder, H. (2004). *Instalação de ar condicionado* (6th ed.). LTC Editora.
- Data Science Academy. (2019). *Deep Learning Book*. <http://deeplearningbook.com.br/o-que-e-aprendizagem-por-reforco/>
- Filho, O., Araujo, S., & Florio, A. (2006). *Detecção automática do aquecimento em componentes de um sistema industrial , baseada no reconhecimento de imagens térmicas* . 1–9.
- Inficon. (2019a). (29) *Leak Testing Methods Part 1: Vacuum Chamber Leak Testing Inside-Out*. You Tube. <https://www.youtube.com/watch?v=i7JOCsEugGY>
- Inficon. (2019b). (29) *Leak Testing Methods Part 2: Sniffing*. You Tube. <https://www.youtube.com/watch?v=HAJaVrZSOJI>
- International Energy Agency. (2018). *The Future of Cooling: Opportunities for energy-efficient air conditioning*. [www.iea.org](http://www.iea.org)
- Ishikawa, K. (1990). *Introduction to quality control* (1st ed.). 3A Corporation.
- Jadin, M. S., & Ghazali, K. H. (2014). Gas leakage detection using thermal imaging technique.

*Proceedings - UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, UKSim 2014*, 302–306. <https://doi.org/10.1109/UKSim.2014.95>

Kohavi, R.; Provost, F. (1998). *Glossary of Terms*. 271–274. <https://doi.org/https://doi.org/10.1023/A:1017181826899>

Langley, G., Moen, R., Nolan, K., Nolan, T., Norman, C., & Provost, L. (2009). *The Improvement Guide: A Practical Approach to Enhancing Organizational Performance (2nd Edition)* (2nd ed.). Jossey-Bass.

LARHUD. (2018). *Orange*. LARHUD Laboratório Em Rede de Humanidade Digital. <http://www.larhud.ibict.br/index.php?title=Orange>

Lin, M. (2018). *Data science cheatsheet*. <http://mavericklin.com>

MARQUES, N., & Araujo, I. (2009). *TEXTOS DE APOIO AO PROFESSOR DE FÍSICA – IF – UFRGS – MARQUES, N. L. R. & ARAUJO, I. S. – v.20 nº 5* (Vol. 20, p. 73). Universidade Federal do Rio Grande do Sul.

Matos, P. F., Lombardi, L. de O., Ciferri, R. R., Pardo, T. A. S., Ciferri, C. D. de A., & Vieira, M. T. P. (2009). *Relatório Técnico “Métricas de Avaliação.”* <http://conteudo.icmc.usp.br/pessoas/taspardo/TechReportUFSCar2009a-MatosEtAl.pdf>

Moen, R., Nolan, T., & Provost, L. (2012). *Quality Improvement through Planned Experimentation, 3rd edition* (3rd ed.). Mc Graw Hill. <https://doi.org/10.1080/00224065.2015.11918108>

Moen, R., & Norman, C. (n.d.). *Evolution of the PDCA Cycle*.

Mohammed, M., Khan, M. B., & Bashie, E. B. M. (2016). Machine learning: Algorithms and applications. In *Machine Learning: Algorithms and Applications* (Issue July). <https://doi.org/10.1201/9781315371658>

Montgomery, D. C., & Wiley, J. (2013). *Design and Analysis of Experiments Eighth Edition*. [www.wiley.com/go/permissions](http://www.wiley.com/go/permissions).

Oliveira, R. A. de, & Silva, A. P. B. da. (2014). William Herschel, os raios invisíveis e as primeiras ideias sobre radiação infravermelha. *Revista Brasileira de Ensino de Física*, 36(4). <https://doi.org/10.1590/s1806-11172014000400022>

- PANDE, P. S., NEUMAN, R. P., & CAVANAGH, R. R. (2000). *The six sigma way: How GE, Motorola, and Other Top Companies Are Honing Their Performance* (Primeira). McGraw-Hill. <https://doi.org/10.1036/0071358064>
- Paramenter, D. (2007). *KEY PERFORMANCE INDICATORS - Developing, Implementing, and Using Winning KPIs*.
- Reis, E. A. (2002). *Análise Descritiva de Dados*. [www.est.ufmg.br](http://www.est.ufmg.br)
- Ribeiro, J. M. (2017). *O conceito da indústria 4.0 na confecção: análise e implementação*. 1–99. <http://repositorium.sdum.uminho.pt/handle/1822/49413>
- Rottländer, H., Umrath, W., & Voss, G. (2016). *Fundamentals of leak detection* (p. 48). Leybold. [www.leybold.com](http://www.leybold.com)
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3, n 3, 210–229. <https://doi.org/10.1147/rd.33.0210>
- SUFRAMA. (n.d.). *Polo Industrial de Manaus*. Retrieved July 17, 2019, from <http://www.suframa.gov.br/invest/zona-franca-de-manaus-pim.cfm>
- Tom Mitchell. (1997). *Machine Learning textbook*. McGraw Hill.
- Vinogradov, M. L., Kostrin, D. K., Karganov, M. V., & Tiskovich, V. Y. (2016). How to choose a leak detection method. *Proceedings of the 2016 IEEE North West Russia Section Young Researchers in Electrical and Electronic Engineering Conference, EIConRusNW 2016, 1*, 100–104. <https://doi.org/10.1109/EIConRusNW.2016.7448130>
- Vollmer, Michael; Mollmann, K.-P. (2010). *Infrared Thermal Imaging*. Wiley-VCH.
- Web Arcondicionado. (2019). *A História do Ar-Condicionado: Linha do Tempo Completa*. Portal WebArCondicionado. <https://www.webarcondicionado.com.br/a-historia-do-ar-condicionado>
- Werkema, M. C. C. (1995). *As ferramentas da qualidade no gerenciamento de processos* (1a ed.). Editora de Desenvolvimento Gerencial.