

**UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ENGENHARIA DE SÃO CARLOS**

**Rafael Prux Fehlberg**

**Uma Nova Abordagem para Estimação de Grau de  
Polimerização em Transformadores de Força utilizando  
Aprendizagem de Máquina.**

**São Carlos**

**2023**



**Rafael Prux Fehlberg**

**Uma Nova Abordagem para Estimação de Grau de Polimerização em Transformadores de Força utilizando Aprendizagem de Máquina.**

Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, para obtenção do título de Mestre em Ciências - Programa de Pós-Graduação em Engenharia Elétrica.

Área de concentração: Sistemas Dinâmicos

Orientador: Prof. Dr. Rogério Andrade Flauzino

Coorientador: Prof. Dr. Ivan Nunes da Silva

**Trata-se da versão corrigida da dissertação. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.**

**São Carlos**

**2023**

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,  
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS  
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da  
EESC/USP com os dados inseridos pelo(a) autor(a).

Prux Fehlberg, Rafael  
Uma Nova Abordagem para Estimação de Grau de  
Polimerização em Transformadores de Força utilizando  
Aprendizagem de Máquina / Rafael Prux Fehlberg;  
orientador Rogério Andrade Flauzino; coorientador Ivan  
Nunes da Silva. São Carlos, 2023.

Dissertação (Mestrado) - Programa de  
Pós-Graduação em Engenharia Elétrica e Área de  
Concentração em Sistemas Dinâmicos -- Escola de  
Engenharia de São Carlos da Universidade de São Paulo,  
2023.

1. Envelhecimento. 2. Transformador de Potência.  
3. Ativos. 4. Grau de Polimerização. 5. Papel. I.  
Título.

## FOLHA DE JULGAMENTO

Candidato: Engenheiro **RAFAEL PRUX FEHLBERG**.

Título da dissertação: "Uma nova abordagem para estimação de grau de polimerização em transformadores de força utilizando aprendizagem de máquina".

Data da defesa: 13/09/2023.

### Comissão Julgadora

### Resultado

Prof. Associado **Rogério Andrade Flauzino**  
**(Orientador)**

(Escola de Engenharia de São Carlos – EESC/USP)

Aprovado

Profa. Dra. **Lucíola Campestrini**

(Universidade Federal do Rio Grande do Sul/UFRGS)

Aprovado

Prof. Dr. **Eduardo Coelho Marques da Costa**

(Escola Politécnica/EP-USP)

Aprovado

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:

Prof. Associado **Marcelo Andrade da Costa Vieira**

Presidente da Comissão de Pós-Graduação:

Prof. Titular **Carlos De Marqui Junior**



Dedico este trabalho aos meus pais, Lizeth e Arno, que sempre me apoiaram em todas as minhas decisões e me deram o suporte necessário para alcançar meus objetivos. Sem o amor, o incentivo e o exemplo de perseverança que eles me proporcionaram, eu não estaria aqui hoje.

Também dedico esta dissertação à minha esposa, Gabriele, que sempre esteve ao meu lado, mesmo nos momentos mais difíceis. Obrigado por me apoiar e me motivar em todos os momentos. Aos meus filhos, Martina e Santiago, pela ausência durante todo o período de estudos.

Por fim, dedico este trabalho a todas as pessoas que, de alguma forma, contribuíram para minha formação acadêmica e pessoal. Seja por meio de uma palavra de encorajamento ou simplesmente por me inspirarem a seguir em frente, todos foram importantes para minha trajetória.

Espero que este trabalho possa ser útil para a comunidade acadêmica e para a sociedade em geral. Mais uma vez, agradeço a todos que fizeram parte desta jornada.



## **AGRADECIMENTOS**

Primeiramente, gostaria de expressar minha profunda gratidão a Dr. Rogério Flauzino, pela orientação, apoio e incentivo ao longo desta jornada acadêmica. Suas valiosas sugestões e críticas construtivas foram fundamentais para o sucesso desta dissertação.

À Neoenergia e à ANEEL pelos recursos disponibilizados ao projeto de Pesquisa e Desenvolvimento: Metodologia Multicritérios e Software Inteligente para Planejamento Otimizado de Manutenção (PD-05160-1804/2018).

Aos colegas Daniel Carrijo Polônio Araujo, Gabriel de Souza Pereira Gomes e Murilo Marques que ajudaram a tornar este trabalho em uma agradável jornada de aprendizagem.

Não posso deixar de agradecer minha família e amigos que me apoiaram e encorajaram durante todo o processo. Seu amor e apoio incondicional foram essenciais para minha motivação e perseverança.

Por fim, expresso minha gratidão a todos aqueles que, direta ou indiretamente, contribuíram para a realização deste trabalho. Espero que os resultados apresentados nesta dissertação possam contribuir de alguma forma para a área de Gestão de Ativos e Manutenção de Transformadores.



*"Todos os modelos estão errados, mas alguns são úteis"*  
*George Box*



## RESUMO

Fehlberg, R.P. **Uma Nova Abordagem para Estimção de Grau de Polimerização em Transformadores de Força utilizando Aprendizagem de Máquina.** 2023. 58p. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2023.

O principal indicador utilizado para avaliar o estado da isolação s3lida em equipamentos de pot4ncia atualmente 4 o grau de polimeriza33o (GP), medido por um ensaio que necessita de uma amostra de papel da parte ativa. Sendo este um ensaio invasivo, outros subprodutos do envelhecimento s3o utilizados para estimar o GP, como o 2-fal. Este trabalho apresenta um estudo sistem3tico onde t4cnicas de *machine learning* s3o utilizadas para estimar o GP a partir do 2-fal, da rela33o CO<sub>2</sub>/CO e de outros indicadores. Os resultados s3o promissores, indicando que 2-fal, CO<sub>2</sub>/CO, a f3rmula de Chendong e a pot4ncia do equipamento podem ser utilizados em conjunto para melhor previs3o do valor atual do GP do papel, logo, da sua vida 3til.

**Palavras-chave:** Envelhecimento, Transformador de Pot4ncia, Ativos, Grau de Polimeriza33o, Papel.



## ABSTRACT

Fehlberg, R.P. **New Approach to estimate DP on Power Transformer using Machine Learning**. 2023. 58p. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2023.

The main indicator used to assess the condition of solid insulation in power equipment currently is the degree of polymerization (DP), measured by a test that requires a sample of the active part paper. As this is an invasive test, other aging byproducts are used to estimate the DP, such as 2-fal. This work presents a systematic study where machine learning techniques are used to estimate DP from 2-fal, CO<sub>2</sub>/CO ratio, and other indicators. The results are promising, indicating that 2-fal, CO<sub>2</sub>/CO, the Chendong formula, and equipment power can be used together to better predict the current value of paper DP, and hence its service life.

**Keywords:** Aging, Power Transformer, Assets, Degree of Polymerization, Paper.



## LISTA DE FIGURAS

Figura 1 – Vizinho mais próximo - k . . . . .	33
Figura 2 – Representação de uma Rede Bayesiana . . . . .	34
Figura 3 – Exemplo de árvore de decisão . . . . .	35
Figura 4 – Representação de uma máquina de vetores de suporte . . . . .	36
Figura 5 – Metodologia utilizada . . . . .	40
Figura 6 – Histograma dos dados . . . . .	41
Figura 7 – Relação entre GP e 2-fal . . . . .	41
Figura 8 – Matriz de correlação de Spearman . . . . .	42
Figura 9 – Histograma dos atributos . . . . .	43
Figura 10 – Relação GP e 2-fal . . . . .	44
Figura 11 – Relação GP e $CO_2/CO$ . . . . .	44
Figura 12 – Relação entre 2-fal, $CO_2/CO$ e GP . . . . .	45
Figura 13 – Relação entre GP predito e GP medido para o treinamento . . . . .	48
Figura 14 – Relação entre GP predito e GP medido para o teste . . . . .	48
Figura 15 – Relação entre GP predito e GP medido para o treinamento . . . . .	49
Figura 16 – Relação entre GP predito e GP medido para o teste . . . . .	50
Figura 17 – Relação entre GP predito e GP medido para o treinamento . . . . .	51
Figura 18 – Relação entre GP predito e GP medido para o teste . . . . .	51
Figura 19 – Relação entre GP real e GP Chendong . . . . .	52
Figura 20 – Relação entre GP predito e GP medido para o treinamento . . . . .	52
Figura 21 – Relação entre GP predito e GP medido para o teste . . . . .	53
Figura 22 – Erro modelos . . . . .	54



## LISTA DE TABELAS

Tabela 1 – Dados extraídos de (ASSUNÇÃO, 2007) . . . . .	38
Tabela 2 – Dados de referência dos Transformadores . . . . .	39
Tabela 3 – Parâmetros estatísticos dos dados . . . . .	40
Tabela 4 – Descrição dos dados . . . . .	42
Tabela 5 – Matriz de Correlação . . . . .	45
Tabela 6 – <i>Mean Average Error</i> ao longo dos <i>folds</i> para cada grau de polinômio . . . . .	47
Tabela 7 – Métricas de avaliação do modelo utilizando os atributos: 2-fal e a relação CO <sub>2</sub> /CO . . . . .	49
Tabela 8 – Métricas de avaliação do modelo utilizando os atributos: 2-fal, relação CO <sub>2</sub> /CO e GP calculado utilizando a fórmula de Chendong . . . . .	51
Tabela 9 – Métricas de avaliação do modelo utilizando os atributos: 2-fal, relação CO <sub>2</sub> /CO, GP calculado utilizando a fórmula de Chendong e a Potência do Transformador . . . . .	53
Tabela 10 – Comparação entre os modelos utilizados . . . . .	53



## LISTA DE ABREVIATURAS E SIGLAS

2-fal	2-Furfuraldeído
5-MEF	5-Metil 2-Furfural
DGA	Análise de Gás Dissolvido
EESC	Escola de Engenharia de São Carlos
ENN	Rede Neural Estendida
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
GP	Grau de Polimerização
LS	Mínimos Quadrados
MAE	Erro Médio Absoluto
OIP	Papel Impregnado de Óleo
RNA	Rede Neural Artificial



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>23</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>25</b>
2.1	Influências no prazo de vida do papel de isolamento	25
2.2	Métodos para o cálculo de vida útil restante	26
2.2.1	CO <sub>2</sub> e CO	26
2.2.2	Compostos Furânicos	27
2.3	Regressão Ridge	28
2.4	Monitoramento on-line do sistema de isolamento	30
2.5	Teoria de <i>Shapley additive explanations</i> (SHAP)	31
2.6	Detecção de defeitos com Inteligência Artificial	31
2.6.1	Aprendizado supervisionado	32
<b>3</b>	<b>METODOLOGIA</b>	<b>37</b>
3.1	Visão geral do problema	37
3.1.1	Descrição dos dados	38
3.1.2	Abordagem utilizada	40
3.1.3	Análise exploratória e tratamento inicial dos dados	40
3.1.3.1	Dados de (ASSUNÇÃO, 2007)	40
3.1.3.2	Dados de (TEYMOURI; VAHIDI, 2017)	42
3.1.4	Conclusões da análise exploratória	46
<b>4</b>	<b>RESULTADOS OBTIDOS</b>	<b>47</b>
4.1	Modelo de previsão do GP a partir do 2-fal	47
4.2	Modelo de previsão do GP a partir do 2-fal e da razão CO <sub>2</sub> /CO	47
4.3	Modelo de previsão do GP a partir do 2-fal, da razão CO <sub>2</sub> /CO e da fórmula de Chendong	50
4.4	Modelo de previsão do GP a partir do 2-fal, da relação CO <sub>2</sub> /CO, da fórmula de Chendong e da potência do transformador	52
4.5	Comparação dos modelos	53
<b>5</b>	<b>CONCLUSÃO</b>	<b>55</b>
	<b>REFERÊNCIAS</b>	<b>57</b>



## 1 INTRODUÇÃO

Para a tomada de decisões acerca da manutenção e operação de transformadores de potência, um dos principais fatores é a vida útil do ativo. A vida útil da isolação é uma das grandezas mais importantes dos transformadores, pois é relacionada diretamente com a vida útil do ativo.

Estudos têm sido desenvolvidos de modo a estimar a vida útil da isolação de transformadores, como (ELELE et al., 2022), (SHUTENKO; PONOMARENKO, 2022) e (ORIA et al., 2019). Um dos métodos mais aceitos na atualidade para a determinação da vida útil é a correlação com o nível do grau de polimerização da celulose da isolação. O grau de polimerização (GP) está diretamente ligado ao envelhecimento do papel isolante do transformador; quanto menor seu valor, mais velha está a isolação. De acordo com artigos científicos publicados atualmente e com fórmulas derivadas destes estudos, é possível estimar a vida restante em anos pelo valor do GP. Ainda assim, o ensaio do GP do papel pode ser custoso e de alta dificuldade, uma vez que é necessário acessar o núcleo do transformador e retirar uma pequena parte do papel isolante para análise.

Por conta da dificuldade da realização de ensaios do GP, outros estudos conseguiram com sucesso relacionar o grau de polimerização da isolação com o nível de compostos furânicos presentes no óleo isolante. Entre os compostos furânicos existentes no líquido isolante, o composto 2-furfuraldeído (2-fal) é o mais utilizado para realizar essa correspondência com o GP e, através desse, a estimativa de vida útil.

Esse método é utilizado amplamente por concessionárias e fabricantes de transformadores, uma vez que, para a realização desse tipo de ensaio, é necessário apenas retirar uma amostra de óleo do transformador, o que já é realizado costumeiramente para as análises cromatográficas e físico-químicas do óleo isolante.

Assim como a utilização dos compostos furânicos, alguns trabalhos também correlacionam a proporção dos gases  $\text{CO}_2$  e  $\text{CO}$ , medida através de cromatografia gasosa, com o tempo de vida útil do transformador. E outros estudos já tentaram abordar o assunto, porém sempre considerando técnicas separadas, como em (ACIU et al., 2021). Tendo como base esses estudos já consolidados, neste trabalho é apresentada uma proposta para unificação dos métodos como uma nova abordagem para a estimação do grau de polimerização e, portanto, da vida útil de transformadores, utilizando a aprendizagem de máquinas para modelamento do comportamento do transformador, objetivando uma previsão mais precisa do tempo de vida.

A unificação dos métodos visa, primeiramente, a realização da correlação entre os métodos de estimação já citados. É importante destacar também que essa correlação entre

os métodos é necessária para que haja uma melhor assertividade no resultado da análise, ou seja, da vida útil do transformador. Por fim, o emprego de técnicas de aprendizagem de máquinas é necessário para consolidação do modelo do transformador.

De modo sintetizado, este trabalho visa apresentar um modelo de auxílio aos especialistas para tomada de decisão acerca da vida útil de transformadores.

## 2 REVISÃO BIBLIOGRÁFICA

Uma vez que os transformadores de potência são componentes valiosos das redes de energia elétrica, a determinação confiável da vida útil restante de um transformador é muito importante para as concessionárias de energia. Transformadores consistem basicamente de núcleo, fio, papel (celulose) e óleo. Assim, seu desempenho e vida útil remanescente dependem principalmente do estado do papel impregnado de óleo. A troca de óleo durante a vida útil de um transformador é possível, mas não é possível trocar o papel de isolamento. Por isso, o prazo de vida de um transformador é sempre o prazo de vida da sua isolação de papel (TEYMOURI; VAHIDI, 2017).

### 2.1 Influências no prazo de vida do papel de isolamento

Em um transformador de potência, as partes energizadas são separadas por um isolamento, que geralmente consiste em papel isolante impregnado com óleo, também isolante. Durante a operação, ocorre o processo de degradação ao longo do tempo, em que a capacidade do papel isolante de resistir às tensões mecânicas que ocorrem durante as correntes internas e externas de curto-circuito diminui significativamente, colocando uma ameaça sobre a operação estável do equipamento e da energia no sistema. De acordo com as estatísticas, falhas do transformador causadas pelo envelhecimento do isolamento interno são responsáveis por uma parte significativa dos acidentes com transformadores. Portanto, é essencial investigar as características de envelhecimento do sistema de isolamento com o objetivo de uma avaliação de vida melhorada (MÜNSTER et al., 2018).

O envelhecimento de papel isolante, que consiste de celulose, é caracterizado pelo grau de polimerização (GP), o número médio de monômeros de glicose nas cadeias poliméricas de celulose (MÜNSTER et al., 2018). Normalmente, considera-se que o fim da vida útil dos transformadores é alcançado quando o GP do papel cai para um valor de 200, e sua resistência à tração é de aproximadamente 40 % de seu valor inicial. Essa opinião é baseada na suposição de que a força de tração é, então, muito baixa para resistir às vibrações mecânicas e forças eletromecânicas produzidas por curtos-circuitos externos em transformadores. No entanto, essa suposição não é suportada por observações em transformadores (DUVAL et al., 2017).

Isso sugere que a falha dielétrica do isolamento de óleo/papel degradado é principalmente devido ao óleo em vez do papel, e reforça a importância crítica de manter a boa qualidade do óleo em toda a vida dos transformadores, principalmente em seus últimos anos, recuperando o óleo sempre que necessário. Um transformador que experimenta um envelhecimento normal, sem deficiências dielétricas ou elétricas, não está no final de sua vida útil quando atinge um GP baixo calculado ou medido. No entanto, para

continuar operando normalmente, é imprescindível que não haja nenhuma indicação de deficiência elétrica ou dielétrica dentro dele ou de seus componentes, por exemplo, buchas e comutadores em carga, e que a rede seja estável (DUVAL et al., 2017).

Portanto, com base nas observações apresentadas por (DUVAL et al., 2017), sugere-se que transformadores com um GP de papel de 200, incluindo transformadores operando em altas tensões e classificações de potência (MVA), podem ser mantidos em serviço por mais vários anos, talvez até que o GP caia para 150 ou 100, sem aumentar significativamente o risco de falha. Além disso, parece que a má qualidade de óleo no papel/óleo de isolamento degradado são fatores mais importantes para levar à falha do transformador do que baixo GP do papel. Portanto, é fundamental manter o óleo em boas condições ao longo da vida de um transformador, recuperando o óleo sempre que necessário e possível.

De fato, durante uma operação de longo prazo, o isolamento do papel e óleo degrada-se gradualmente sob os efeitos sinérgicos de energia térmica, elétrica e fatores químicos (LIN et al., 2015). Recentes observações experimentais revelam que a umidade no isolamento em combinação com temperatura e estresse elétrico tende a envelhecer o isolamento muito mais rapidamente em comparação com o isolamento seco, e é responsável por falhas prematuras no papel impregnado de óleo (OIP). Na verdade, a umidade foi identificada como um dos principais agentes de degradação do isolamento. Em vista dessas evidências convincentes contra o papel deletério desempenhado pela umidade no envelhecimento catalítico de OIP, surgiu a necessidade de desenvolver modelos modificados de falha de isolamento de tensão combinada envolvendo a umidade como um agente competitivo (MANDLIK; RAMU, 2014).

## 2.2 Métodos para o cálculo de vida útil restante

Uma vez que o GP é medido por meio de um procedimento de teste destrutivo no qual o transformador é desmontado, este não pode ser utilizado em uma situação real. Portanto, torna-se necessário identificar certos parâmetros de diagnóstico, que são funções da quantidade de envelhecimento, para substituir o GP como um índice de falha. Em particular, 2-furfuraldeído (2-fal) e óxidos de carbono, CO e CO<sub>2</sub>, liberados devido à decomposição térmica da celulose, são indicadores da quantidade de envelhecimento. Quando essas quantidades atingem os valores limites prescritos, o isolamento é considerado como tendo falhado (MANDLIK; RAMU, 2014).

### 2.2.1 CO<sub>2</sub> e CO

Os processos de envelhecimento do papel geram vários subprodutos em transformadores, alguns dos quais são estáveis, enquanto outros desaparecem com o passar do tempo. Se qualquer falha incipiente ocorrer devido a problemas elétricos ou tensões térmicas, pode-se ter a decomposição do óleo do transformador, do cartão prensado e

de outros materiais isolantes e, em seguida, gerar gases de subproduto, como hidrogênio ( $H_2$ ), metano ( $CH_4$ ), acetileno ( $C_2H_2$ ), etileno ( $C_2H_4$ ), etano ( $C_2H_6$ ), monóxido de carbono (CO) e dióxido de carbono ( $CO_2$ ), alguns dos quais serão dissolvidos no óleo. A taxa de geração aumenta exponencialmente com o aumento da temperatura. CO e  $CO_2$  também são gerados em baixa concentração através da oxidação do óleo, e suas concentrações no óleo são comumente medidas usando análise de gás. A norma IEC 60599 especifica a faixa de  $3 < CO_2/CO < 10$  para envelhecimento normal do papel, onde  $CO_2/CO$  é a proporção das concentrações de gás. Uma proporção  $< 3$  indica uma falha envolvendo o isolamento do papel (TEYMOURI; VAHIDI, 2017).

A relação  $CO_2/CO$  também pode ser usada para estimar a condição do papel de isolamento. As principais vantagens do uso da razão  $CO_2/CO$ , em vez da concentração de 2-fal, são que (TEYMOURI; VAHIDI, 2017):

- a) a proporção de  $CO_2/CO$  é muito mais estável;
- b) se houver vazamento de gás da atmosfera no transformador (entrada de ar), a proporção não irá variar tanto quanto as concentrações individuais de  $CO_2$  e CO.

A análise de gás dissolvido (DGA) é uma das mais eficazes metodologias para monitoramento de condições de transformadores de energia e é amplamente utilizado para detectar as falhas incipientes.

### 2.2.2 Compostos Furânicos

Compostos furânicos (ou furanos) também são gerados pela degradação do papel isolante impregnado com óleo e são consideravelmente solúveis em óleo. Furanos incluem 2-fal, 5-metil 2-furfural (MEF), 2-furfurilálcool, 2-acetilfurano e 5-hidroximetil-2-furfural. A produção deles é mencionada na norma IEC 61198. Altas concentrações ou aumentos em suas concentrações geralmente indicam degradação do papel devido ao envelhecimento ou condições de falha. Os furanos dissolvidos em óleo podem causar a deterioração do papel em temperaturas inferiores a aproximadamente  $100\text{ }^\circ\text{C}$ . Em tais temperaturas, furanos são gerados principalmente por meio da destruição de hemicelulose, que compõe cerca de 6 a 7 % do peso do papel e é mecanicamente muito mais fraca do que a celulose. Furanos também são gerados no óleo por meio da degradação da celulose em temperaturas na faixa de  $100$  a  $200\text{ }^\circ\text{C}$  (TEYMOURI; VAHIDI, 2017).

A medição das concentrações de furano no óleo é um método muito mais fácil de estimar o GP do papel do que medir diretamente em amostras de papel. 2-fal é preferido como um indicador químico do valor de GP porque é mais estável do que outros furanos em temperaturas abaixo de  $140\text{ }^\circ\text{C}$ . A concentração de 2-fal é sempre maior do que de outros furanos, e este é gerado principalmente quando o valor do GP cai abaixo de 400. Um aumento nas concentrações de umidade e oxigênio no transformador aumentam a

taxa de geração de 2-fal. Sugere-se que a concentração de furanos idealmente deve ser inferior a 0,1 ppm no óleo de um transformador novo, aumentando, mas não excedendo, 1 ppm durante a vida útil do transformador. Contudo, concentrações de até 10 ppm foram relatadas em muitos transformadores em atividade (TEYMOURI; VAHIDI, 2017).

Pelo menos quatro relações entre a concentração de 2-fal, F (em mg/L ou ppm) e o valor GP foram propostas. Essas quatro relações são baseadas inteiramente em dados experimentais ao invés da teoria científica, e dependem de fatores como a massa de papel, as concentrações de umidade e ácido no óleo, e condições de carregamento. Portanto, pode-se dizer que uma relação quantitativa rigorosamente derivada entre a concentração de 2-fal e valor do GP ainda não foram estabelecidas (TEYMOURI; VAHIDI, 2017).

Os principais problemas com o uso da concentração de 2-fal para estimar o GP do papel são (TEYMOURI; VAHIDI, 2017):

- a) 2-fal é gerado em diferentes formas e, portanto, não é exclusivamente o resultado da degradação de celulose;
- b) 2-fal não é estável em temperaturas acima 110 °C;
- c) a presença de oxigênio, água e ácidos afeta a concentração de 2-fal.

Conclui-se que as previsões da vida útil restante de um transformador baseado exclusivamente na concentração de 2-fal no óleo podem nem sempre ser confiáveis.

### 2.3 Regressão Ridge

Um dos tópicos de pesquisa mais interessantes durante os anos setenta e oitenta foi uma nova proposta de método, chamado de regressão Ridge, para estimar os coeficientes de regressão linear múltipla. A popularidade deste tópico pode ser atribuída à importância do problema que ele aborda - colinearidade no contexto de regressão linear múltipla - e a facilidade da adequação da metodologia a ser implementada na prática com base em um exame do chamado Ridge *trace*, um gráfico da regressão estimada de coeficientes em função do parâmetro Ridge.

O impacto negativo da colinearidade no estimador de mínimos quadrados (LS) em um contexto de regressão é conhecido. As abordagens para mitigar esse impacto foram desenvolvidas e muito centradas na eliminação variável, ou seja, na remoção de uma ou mais das variáveis para melhorar o condicionamento da matriz de correlação resultante das demais variáveis independentes. A Regressão Ridge, por outro lado, fornece um meio de resolver o problema de colinearidade sem remover variáveis do conjunto original de variáveis independentes. Isso provou ser um recurso muito atraente em algumas aplicações.

Uma abordagem LS pode fornecer uma base adequada para a previsão geral, mas, quando as variáveis explicativas são não ortogonais, frequentemente falha em dar o

peso adequado para variáveis explicativas individuais usadas como preditores. Em muitos problemas nos quais os dados não são obtidos de um experimento bem projetado ou controlado, estimativas de coeficientes de regressão tendem a se tornar muito grandes em valores absolutos, e é possível que alguns terão até o sinal errado. As chances de encontrar tais dificuldades aumentam quanto mais a previsão de vetores desvia da ortogonalidade (MCDONALD, 2009).

Modelo de regressão Ridge é semelhante à estimativa de mínimos quadrados, a não ser pelo fato de que a variável Ridge de coeficiente é adicionada e otimizada. Isto é, espera-se que o efeito líquido forneça uma estimativa mais confiável. Os coeficientes de Ridge são usados para reduzir uma soma residual de quadrados penalizada (LEAUPRASERT et al., 2020).

Muitos autores utilizam a abordagem de decomposição canônica para problemas de regressão. Isso é muito útil quando são consideradas técnicas de Shrinkage e Ridge para estimativa de parâmetros. Shrinkage e Ridge realmente são ambos métodos de encolhimento em variáveis canônicas. A abordagem canônica fornece percepção da natureza da regressão Ridge e leva a generalizações significativas. Uma dessas generalizações é uma família de dois parâmetros da forma:

$$\hat{\beta}(k, q) = [\mathbf{x}'\mathbf{x} + k(\mathbf{x}'\mathbf{x})^q]^{-1} \mathbf{x}'\mathbf{y}. \quad (1)$$

Regressão Ridge, e o traço Ridge associado, fornecem ao analista uma maneira sistemática de digitalizar uma grande classe de estimadores enviesados. Fazendo isso, frequentemente leva o analista a estimativas significativas de coeficientes de regressão que podem ser usados na prática para avaliar mudanças da variável dependente em relação às variáveis independentes e também preveem valores de resposta em pontos especificados no espaço designado.

Grande parte da literatura sobre regressão Ridge é dedicada a identificar um parâmetro Ridge ideal,  $k$ , para ser usado na prática - se não for ideal, pelo menos um valor  $k$  que irá garantir que o estimador Ridge tenha menor erro médio quadrático do que o estimador LS. Embora existam muitos métodos bons para escolher um valor  $k$  "ideal", não se verificou nenhum vencedor.

Em resumo, a regressão Ridge tem bases fortes para consideração. Existem muitas abordagens sobre a escolha adequada do parâmetro. A escolha final de  $k$  para uma aplicação específica envolvendo variáveis explicativas colineares ainda permanece parte arte e parte ciência. Por mais de três décadas provou-se ser uma ferramenta valiosa para uso por estatísticos aplicados e deve ser explorada rotineiramente em um contexto de regressão múltipla colinear (MCDONALD, 2009).

## 2.4 Monitoramento on-line do sistema de isolamento

A maioria dos transformadores poderá ter uma emergência a partir de uma sobrecarga ocasionalmente, com subsequente perda de vida. No entanto, benefícios substanciais podem ser obtidos operando o transformador além do tempo que se baseia nas classificações da placa de identificação. Há uma necessidade crescente das empresas para empregar transformadores ao máximo, mantendo a confiabilidade do sistema. Atualmente, a carga e a idade dos aparelhos estão aumentando e, portanto, o monitoramento e diagnóstico de transformadores de potência tornam-se cada vez mais importantes, sendo o monitoramento a coleta de dados relevantes durante o serviço (on-line) ou durante manutenção ou períodos de teste (off-line) e, diagnóstico a avaliação técnica e interpretação do dados registrados.

A importância das consequências econômicas do transformador monitorado são a base para a gestão de ativos de transformadores de potência juntamente com a avaliação de risco. Um sistema de monitoramento on-line é particularmente adequado para utilização com transformadores de potência, com o objetivo de garantir um fornecimento confiável de energia elétrica, juntamente com despesas de manutenção reduzidas e um aproveitamento ótimo dos ativos. Além da vida operacional restante do transformador poder ser estimada, registrando dados operacionais importantes, para permitir uma utilização consistente tecnicamente possível de carga do transformador, declarações em relação à capacidade de sobrecarga atual, por exemplo, podem ser feitas. Os sistemas de monitoramento on-line podem ajudar a reduzir custos do ciclo de vida e manutenção de transformadores de potência. A estimativa de vida útil mais precisa leva a uma melhoria nas estratégias de investimento (ARVIND; KHUSHDEEP; DEEPAK, 2008).

A realização de sistemas de diagnóstico on-line consiste em uma composição de elementos adequados para todo o sistema de medição. O sistema de medição digital é, na verdade, a configuração de equipamentos e suas conexões fornecendo uma solução de medição completa. O transformador monitorado é equipado por sensores (Pt100, transformadores de instrumento, analisadores de gás e umidade, etc.) e suas saídas são vinculadas a um ou mais computadores. A função correta e o cumprimento das operações especificadas são fornecidos por um programa com um design adequadamente criado. A possibilidade de modificação flexível de sensores e indicadores digitais de acordo com as necessidades do usuário tem sido a principal vantagem desta solução de aparelhos de medição. O sistema habilita automaticamente o processo de salvamento, visualização e processamento de dados medidos, que eliminam um dos recursos de erros acidentais apresentado pelo registro manual dos valores medidos. Isso é necessário para fornecer a comunicação entre os sensores e o computador, que fornece todas as atividades subsequentes relacionadas com medição. Interface serial de computadores pessoais (RS-232, RS-485), ou os chamados cartões multifuncionais, podem ser usados para esta comunicação. A seleção da interface de comunicação sempre depende dos requisitos do usuário (MENTLIK et al.,

2006).

## 2.5 Teoria de Shapley additive explanations (SHAP)

A seleção das características tem uma enorme influência num modelo de classificação. Acrescentar características que dão mais informação para descrever as amostras tornam a classificação mais fácil. No entanto, mais características implicam uma dimensão superior e um espaço de distribuição de amostras esparso, e pode ocorrer formação insuficiente, levando ao excesso de formação e à maldição da dimensionalidade. Além disso, as variáveis irrelevantes ou repetitivas afetam a importância das características, o que, por sua vez, diminui a precisão da classificação e a sob-reutilização dos recursos informáticos (ZHANG et al., 2022).

SHAP é uma abordagem baseada na teoria dos jogos para descrever o desempenho de um modelo de aprendizagem por máquina. Para produzir um modelo interpretável, SHAP utiliza um método de atribuição de características aditivas, ou seja, um modelo de saída é definido como uma adição linear de variáveis de entrada. O SHAP pode, assim, fornecer boas explicações para modelos locais e globais. Os valores de SHAP podem ser aproximados por vários métodos, tais como *Kernel SHAP*, *Deep SHAP* e *Tree SHAP*. Os valores de interação SHAP garantem coerência às explicações dos efeitos de interação para as previsões individuais. As duas vantagens únicas dos valores SHAP são a sua capacidade de interpretação global e local. Ao contrário das características de importância existentes nos modelos de aprendizagem mecânica, SHAP pode identificar se a contribuição de cada característica de entrada é positiva ou negativa. Além disso, cada observação pode obter o seu valor SHAP. Assim, o SHAP pode ajudar a interpretar o modelo tanto a nível global como local (MANGALATHU; HWANG; JEON, 2020)

## 2.6 Detecção de defeitos com Inteligência Artificial

O objetivo do campo de detecção de defeitos é calcular quanto tempo um equipamento ainda consegue trabalhar em condições confiáveis (KORDESTANI et al., 2021). Então, os defeitos são incertos, faltam medidas para a mensuração das dinâmicas dos defeitos e a maioria deles possui comportamento não-linear, dessa forma, este campo é desafiador.

Neste sentido, existem dois estágios para a detecção de defeitos:

- Modelo de identificação de falhas: através do uso da física ou modelos *black-box* com dados de entrada e saída;
- Predição: a vida útil restante é prevista usando o método de identificação de falhas, podendo ser através de uma via probabilística ou determinística.

Além disso, em (KORDESTANI et al., 2021) são apresentados os principais métodos para se detectar defeitos. Primeiro, existem os métodos baseados em modelos ou em física. Esses métodos podem ser efetivos sempre que existe algum modelo matemático que pode ser utilizado para descrever o processo de falha. Os métodos Kalman são exemplos deste método.

Também existem os métodos baseados em dados ou inteligência artificial. Estes métodos utilizam de técnicas de aprendizagem de máquina, como *Artificial Neural Networks*, *Deep Learning* ou *Adaptative Neural Fuzzy Inference Systems* através de padrões de entradas e saídas para detectar os defeitos, o modelo *black-box* dos dados. Estes métodos são preferíveis quando não se tem muito conhecimento do sistema, devido à habilidade destes métodos de aprender e generalizar relações não-lineares entre entrada e saída de dados. Métodos probabilísticos e baseados em Markov são exemplos destes modelos.

Em terceiro, tem-se os métodos baseados em conhecimento, estes usam o conhecimento profundo de especialistas no sistema. Geralmente estes métodos usam ferramentas como processamento de sinais e lógica Fuzzy. Também existem os sistemas híbridos que combinam vários desses métodos para prever defeitos.

### 2.6.1 Aprendizado supervisionado

A função central do *machine learning* é dizer aos computadores como prever resultados com base em experiências anteriores e esse trabalho é feito por um bom classificador. Classificação é o processo de usar um modelo para prever valores desconhecidos (saída de variáveis), usando vários valores conhecidos.

Sendo assim, na aprendizagem supervisionada, o primeiro passo é lidar com o conjunto de dados. Com o objetivo de realizar um melhor treinamento no conjunto de dados, o especialista poderia sugerir uma melhor seleção de recursos. Se o especialista não estiver ao alcance, então a outra abordagem é "força bruta", o que significa medir tudo disponível na esperança de que os recursos certos (informativos, relevantes) possam ser isolados. No entanto, um conjunto de dados coletados pelo método de "força bruta" não é diretamente adequado para indução. Em última análise, na maioria dos casos, contém ruído e faltam valores de recursos e, portanto, requerem pré-processamento. Na próxima etapa, a preparação de dados é uma função fundamental. Uma série de técnicas foram introduzidas por diferentes pesquisadores para lidar com dados ausentes (IQBAL; YAN, 2015).

Portanto, uma necessidade comum ao analisar conjuntos de dados do mundo real é determinar quais instâncias se destacam como sendo diferentes para todos os outros. Essas instâncias são conhecidas como anomalias e o objetivo da detecção de anomalias é determinar todas essas instâncias de uma forma orientada por dados. Essas anomalias podem ser causadas por erros nos dados, mas às vezes são indicativos de um novo processo

subjacente anteriormente desconhecido (CHALAPATHY; CHAWLA, 2019).

Métodos supervisionados (também conhecidos como métodos de classificação) exigem um conjunto de treinamento rotulado contendo amostras normais e amostras anômalas para construir o modelo preditivo. Teoricamente, os métodos supervisionados fornecem uma melhor detecção do que os métodos semi-supervisionados e não supervisionados, uma vez que eles têm acesso a mais informações. No entanto, existem alguns problemas técnicos, que fazem com que esses métodos não pareçam precisos como deveriam ser. O primeiro problema é a escassez de um conjunto de dados de treinamento que cubra todas as áreas. Além disso, obter rótulos precisos é um desafio e os conjuntos de treinamento geralmente contêm alguns ruídos que resultam em um alarme falso. Os algoritmos supervisionados mais comuns são (OMAR et al., 2013):

- Vizinho mais próximo - k: esta é uma modesta e não paramétrica técnica convencional para classificar amostras. Calcula-se as distâncias aproximadas entre vários pontos nos vetores de entrada e, em seguida, atribui-se o ponto sem classe para a classe de seus vizinhos mais próximos. No processo de criação do classificador vizinho mais próximo - k, "k" é um importante parâmetro e vários valores k podem obter diferentes performances. Se k for muito grande, os vizinhos, que são utilizados na previsão, consumirão muito tempo de classificação e afetarão a precisão da previsão. Este é um dos métodos mais amplamente utilizado para uma classificação de objetos quando há pouco ou nenhum conhecimento prévio sobre a distribuição dos dados. Portanto, é uma boa escolha para realizar análise discriminante quando as densidades de probabilidade são desconhecidas ou difíceis de determinar (MUHAMMAD; YAN, 2015). Na Figura 1 temos a representação desse método.

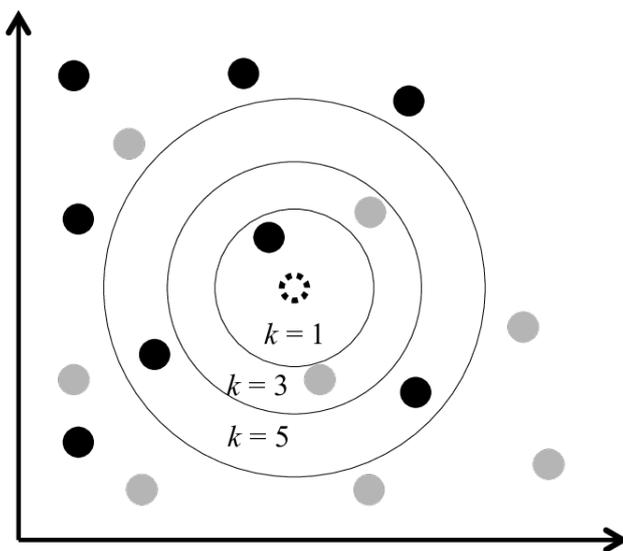


Figura 1: Vizinho mais próximo - k

- Rede Bayesiana: estes são modelos gráficos usados para ilustrar as relações entre eventos ou ideias para inferir probabilidades ou incertezas associadas a essas ideias ou eventos. Recuperação de informações, previsões baseadas em entradas limitadas ou software de reconhecimento são algumas das principais aplicações das Redes Bayesianas (BN) (MUHAMMAD; YAN, 2015). Este é um modelo que codifica relações probabilísticas entre as variáveis de interesse. Esta técnica é geralmente usada para detecção de intrusão em combinação com esquemas estatísticos. Isso tem várias vantagens, incluindo a capacidade de codificação das interdependências entre variáveis e de prever eventos, bem como a capacidade de incorporar conhecimento prévio e dados. Na Figura 2 temos a representação de uma rede Bayesiana.

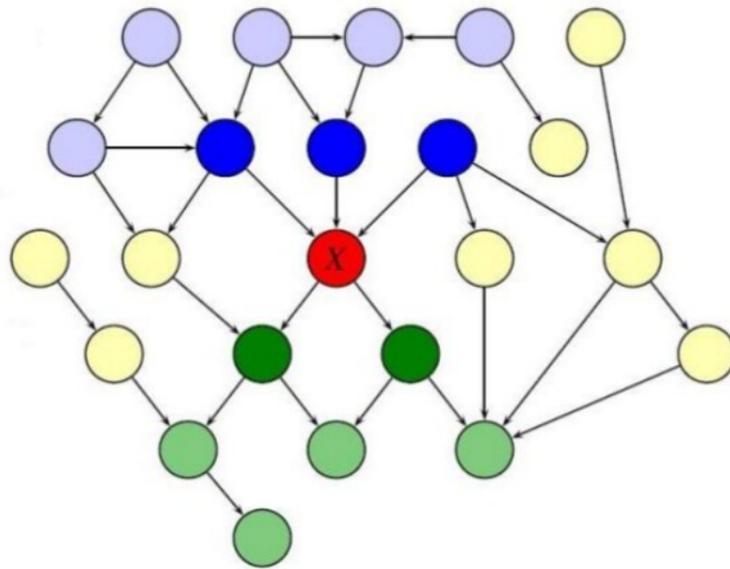


Figura 2: Representação de uma Rede Bayesiana

- Redes neurais supervisionadas: o uso de redes neurais artificiais ganhou popularidade nos últimos anos em reconhecimento de padrões e *machine learning*. A maioria das técnicas populares de *Deep Learning* são construídas a partir de Rede Neural Artificial (RNA). *Deep Learning* pode ser definido como um modelo (por exemplo, redes neurais) com muitas camadas, e treinadas camada à camada (MUHAMMAD; YAN, 2015). Se forem devidamente projetadas e implementadas, as redes neurais têm a capacidade de atender a muitos problemas encontrados por abordagens baseadas em regras. A principal vantagem das redes neurais é sua tolerância a dados imprecisos, informações incertas e sua capacidade de concluir soluções a partir de dados sem ter conhecimento prévio da regularidade nos dados.
- Árvores de decisão: são árvores que classificam as instâncias separando-as com base nos valores de recursos, onde cada nó em uma árvore de decisão representa um recurso em uma instância a ser classificada, e cada ramo representa um valor que o

nó pode assumir. As instâncias são classificadas começando no nó raiz e classificadas com base em seus valores de recursos (MUHAMMAD; YAN, 2015). Uma árvore de decisão é uma árvore que tem três componentes principais: tronco, galhos e folhas. Cada tronco é rotulado com um atributo de característica, que é mais informativo entre os atributos ainda não considerados no caminho da raiz. Cada galho de um tronco é rotulado com um valor do recurso para o recurso do tronco, e cada folha é rotulada com uma categoria ou classe. Uma árvore de decisão pode então ser usada para classificar um ponto de dados começando na raiz da árvore e movendo-se através dele até que uma folha seja alcançada. Um exemplo de árvore de decisão pode ser visto na Figura 3.

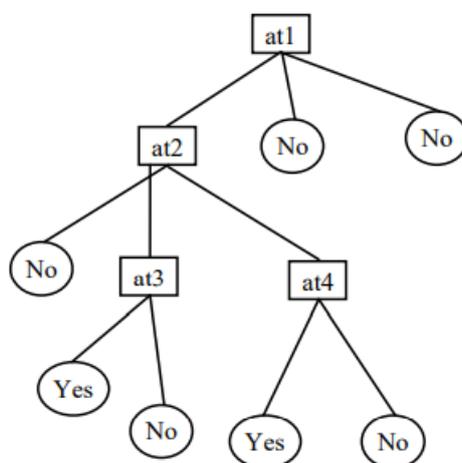


Figura 3: Exemplo de árvore de decisão

- Máquina de vetores de suporte: este é um conjunto de máquinas supervisionadas que foram usadas para classificação, regressão e detecção de *outlier*. Existem vários benefícios para usar esta abordagem, como este método usa um subconjunto de pontos de treinamento na função de decisão (chamados de vetores de suporte), é eficiente em termos de memória, e também é versátil porque contém diferentes funções do kernel que podem ser especificadas para a função de decisão. Kernels comuns são fornecidos, mas também é possível especificar kernels personalizados (MUHAMMAD; YAN, 2015). Nesse método, primeiro mapeia-se o vetor de entrada em uma dimensão superior e, em seguida, obtém-se o hiperplano de separação ideal no espaço de recursos de alta dimensão. Assim, o hiperplano de separação é determinado por vetores de suporte ao invés de todo o treinamento. Portanto, é extremamente robusto para *outliers*. Em particular, esse classificador é projetado para classificação binária. Isto é, separar um conjunto de vetores de treinamento que pertence a duas diferentes classes, que dão suporte às amostras do treinamento perto de um limite de decisão. Também fornece um parâmetro especificado pelo usuário denominado fator de penalidade. Permite que usuários façam uma troca entre o número de amostras

de classificação incorreta e a largura da fronteira de decisão. Na Figura 4 temos sua representação.

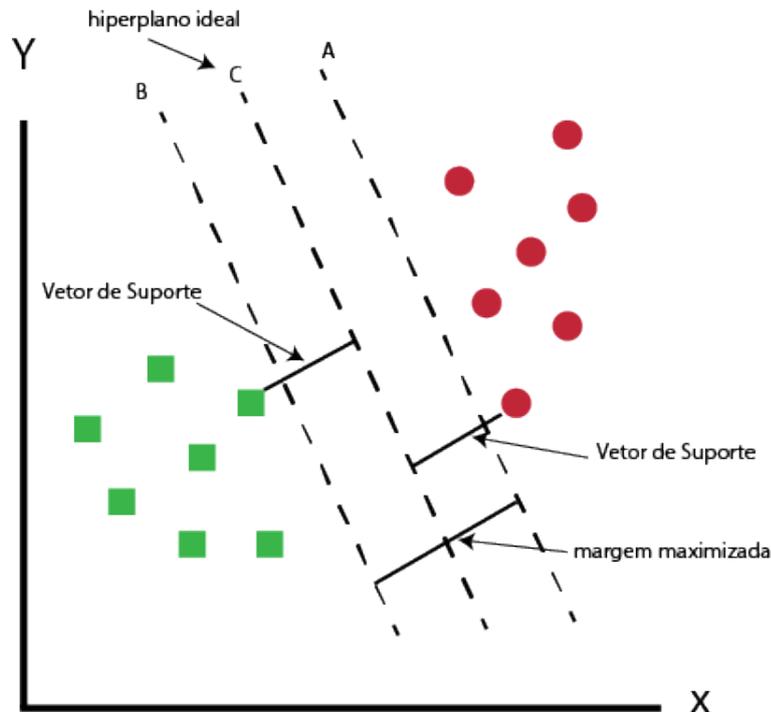


Figura 4: Representação de uma máquina de vetores de suporte

Um dos maiores problemas no campo da pesquisa e desenvolvimento de *Machine Learning* para a rede elétrica é, sem dúvida, a geração de diagnósticos confiáveis sobre as falhas de máquinas. Pensando nesse impasse, diversos pesquisadores e cientistas de dados têm direcionado seus esforços para o desenvolvimento de sistemas com essa capacidade.

Em 2006, os pesquisadores Vilakazi e Marwala ([VILAKAZI; MARWALA, 2006](#)) fizeram uma implementação de Rede Neural Estendida (ENN) para detecção e diagnóstico de falhas em buchas de papel impregnado de óleo (OIP) fabricadas com base nos métodos de taxa de produção IEEE c57.104, IEC 599 e IEEE. O algoritmo implementado foi capaz de gerar diagnósticos sobre a condição da bucha a partir de dados coletados em uma análise dos gases dissolvidos no óleo. Essa análise baseou-se em uma classificação em dois estágios, em que o primeiro estágio verificou se a bucha estava com defeito e o segundo estágio classificou o tipo de falha. Após a implementação, o ENN levou cerca de dois segundos para treinar cada uma das etapas e, de acordo com o relatório da pesquisa, esse método alcançou cerca de 100 % de acurácia de treinamento para ambas as etapas de detecção e classificação.

## 3 METODOLOGIA

### 3.1 Visão geral do problema

A principal causa de falha em transformadores de força é o envelhecimento de sua isolação. Muitos subprodutos são gerados através do aquecimento e resfriamento do óleo em contato com o papel isolante. A decomposição térmica de papel impregnado com óleo gera os gases CO e CO<sub>2</sub>, que são os produtos finais da destruição do papel. A taxa da geração aumenta exponencialmente com o aumento da temperatura do papel. CO e CO<sub>2</sub> também são gerados em baixa concentração através da oxidação do óleo. Compostos furânicos (ou furanos) também são gerados pela degradação do papel isolante impregnado com óleo e são consideravelmente solúveis em óleo. Todos esses compostos podem ser utilizados para avaliar o estado do transformador, auxiliando a evitar falhas devido ao envelhecimento. Porém, seus valores não são indicativos diretos do estado da isolação sólida.

Dessa forma, para indicar o estágio de envelhecimento de transformadores isolados por papel impregnado de óleo de maneira precisa, podemos utilizar a medição do GP (grau de polimerização), que indica quantos monômeros existem em uma cadeia de celulose. Porém, medir o GP de transformadores não é uma tarefa fácil. É necessário que seja removido um pedaço do papel e levado para análise no laboratório. Para remover um pedaço do papel, é necessário desligar o transformador e abri-lo. Essa parada do transformador gera alto custo de manutenção, além do custo da parada em si. Devido a isso, diversos métodos ao longo dos anos foram desenvolvidos para estimar o GP a partir de ensaios mais simples, que não necessitassem da parada do transformador, como o ensaio de cromatografia gasosa e o ensaio físico químico, através da medição de subprodutos do envelhecimento, como gases CO<sub>2</sub> e CO e também 2-furfuraldeído (2-fal), que, conforme mencionado acima, também são indicativos de envelhecimento, por mais que não possuam uma correlação direta.

Nas literaturas anteriores, foram propostas diversas fórmulas para calcular a relação entre a concentração de 2-fal e o valor de GP. Das propostas, podemos citar as de Chendong, Scholnick, De Pablo e Pahlavanpour (TEYMOURI; VAHIDI, 2017), porém, essas fórmulas apresentam valores discrepantes, mostrando que apenas os valores de 2-fal não são suficientes para o cálculo preciso do valor de GP. Dessa forma, esse trabalho propõe uma nova metodologia capaz de calcular o valor de GP utilizando os valores de 2-fal e da relação CO<sub>2</sub>/CO como entradas de um modelo de aprendizagem de máquina. Os passos realizados para esse desenvolvimento são apresentados nas seções a seguir.

## 3.1.1 Descrição dos dados

Neste trabalho, o método proposto é baseado em dados fornecidos por outros trabalhos realizados na mesma área. Podemos dividir os dados em dois grandes grupos, sendo o primeiro deles extraído de (ASSUNÇÃO, 2007), apresentado na Tabela 1. Esses dados possuem diversas medições de 2-fal e as medições de GP correspondentes.

Tabela 1: Dados extraídos de (ASSUNÇÃO, 2007)

2-fal [ppb]	GPM
60	1044
80	971
105	911
136	853
174	797
221,6	744
279,5	693
349,5	644
432	597
531	553
645	511
776	471
925	433
1090	398
1271	365
1472	334
1680	306
1892	279
2107	255
2325	234
2551	214
2784	197
432	597
531	553
645	511
776	471
925	433
1090	398
1271	365
1472	334
1680	306
1892	279
2107	255
2325	234
2551	214
2784	197

O segundo conjunto de dados utilizado foi extraído de (TEYMOURI; VAHIDI, 2017). A Tabela 2 apresenta os dados contendo as medições dos transformadores, como as concentrações dos compostos furânicos, que foram medidos usando cromatografia líquida de alta pressão; a relação  $\text{CO}_2/\text{CO}$  calculada a partir das concentrações de gases  $\text{CO}_2$  e  $\text{CO}$  obtida através de cromatografia gasosa; GP, medido mediante a análise em laboratório; e a potência do transformador (MVA). O segundo conjunto de dados é mais completo que o primeiro e apresenta mais informações que serão consideradas na modelagem do problema.

Tabela 2: Dados de referência dos Transformadores

Transformador	Potência (MVA)	$\text{CO}_2/\text{CO}$	2-fal (ppb)	GPM
1	95	6,2	10	1089
2	105	6,4	13	1042
3	96	6,5	18	983
4	92	6,6	19	971
5	50	6,7	24	915
6	123	6,92	34	858
7	10	6,94	51	790
8	154	6,98	60	776
9	40	7	63	761
10	50	7,1	66	740
11	35	7,21	88	697
12	96	7,32	115	653
13	20	7,36	149	589
14	30	7,5	158	566
15	96	7,7	191	549
16	92	7,74	241	520
17	25	7,78	283	499
18	92	7,8	300	468
19	125	7,85	335	456
20	50	7,88	369	445
21	100	7,9	448	408
22	123	8	544	368
23	50	8,1	596	354
24	96	8,15	637	334
25	100	8,2	740	310
26	115	8,23	844	273
27	92	8,26	958	251
28	85	8,7	1060	243
29	154	8,85	1090	235
30	123	8,9	1180	214

### 3.1.2 Abordagem utilizada

Este trabalho foi desenvolvido conforme o fluxograma apresentado na [Figura 5](#). Primeiramente foi realizada uma análise exploratória e tratamento inicial dos dados obtidos para o modelo de previsão de Grau de Polimerização. Posteriormente foi realizado a aplicação dos dados em cada modelo proposto e avaliados os resultados obtidos de cada modelo, e ao final uma comparação entre os modelos.

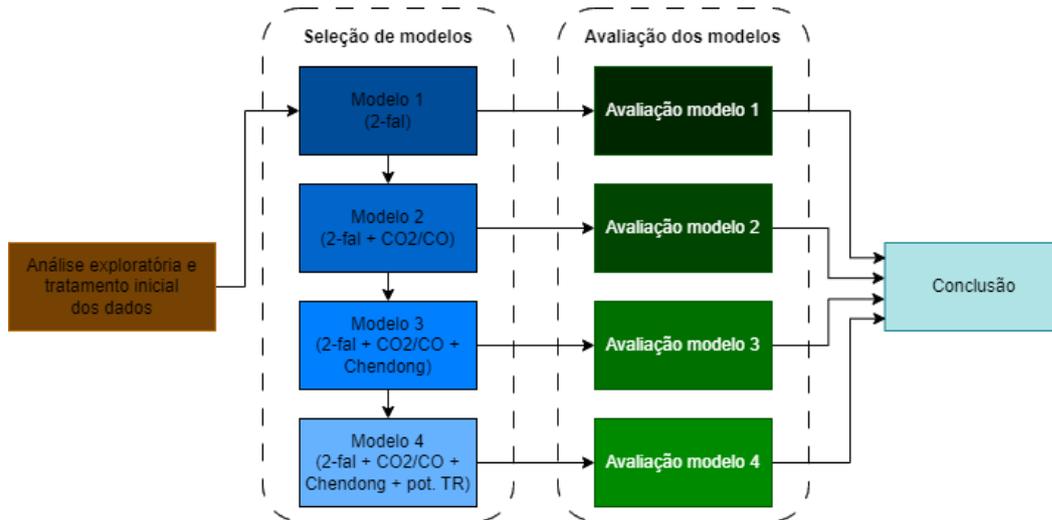


Figura 5: Metodologia utilizada

### 3.1.3 Análise exploratória e tratamento inicial dos dados

A análise exploratória dos dados será realizada separadamente, principalmente devido a dois fatores: os dados têm origem diferente e possuem atributos diferentes. Dessa forma, convém que a análise exploratória seja realizada inicialmente separadamente e, depois, conclusões sejam obtidas a partir da comparação dos resultados.

#### 3.1.3.1 Dados de ([ASSUNÇÃO, 2007](#))

Os dados de ([ASSUNÇÃO, 2007](#)) possuem apenas dois atributos, sendo eles: 2-fal e GP. Inicialmente, é possível verificar os parâmetros estatísticos básicos desses dados, conforme apresentado na [Tabela 3](#).

Tabela 3: Parâmetros estatísticos dos dados

	<b>GPM</b>	<b>2-fal (ppb)</b>
count	36,000000	36,000000
média	470,861111	1176,877778
std	233,966072	863,447395
min	197,000000	60,000000
50 %	415,500000	1007,500000
max	1044,000000	2784,000000

Conforme é possível perceber, metade dos dados estão abaixo de 415 de GP. Somado a isso, a média e a mediana são diferentes, indicando que a distribuição estatística dos dados não se comporta como uma normal. É possível verificar esse fato através de um histograma, conforme mostrado na [Figura 6](#).

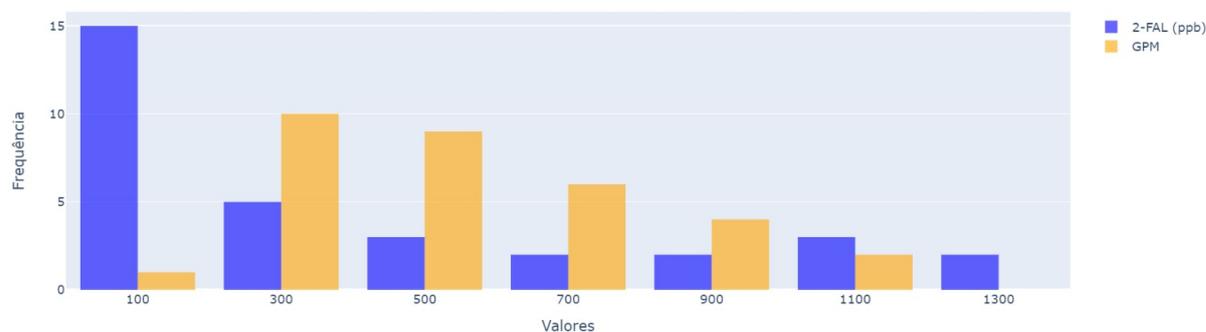


Figura 6: Histograma dos dados

Conforme é possível observar, a maior parte dos dados se concentra para valores de GP e 2-fal baixos, criando um deslocamento à esquerda dos dados. A relação entre esses dois atributos pode ser verificada na [Figura 7](#).

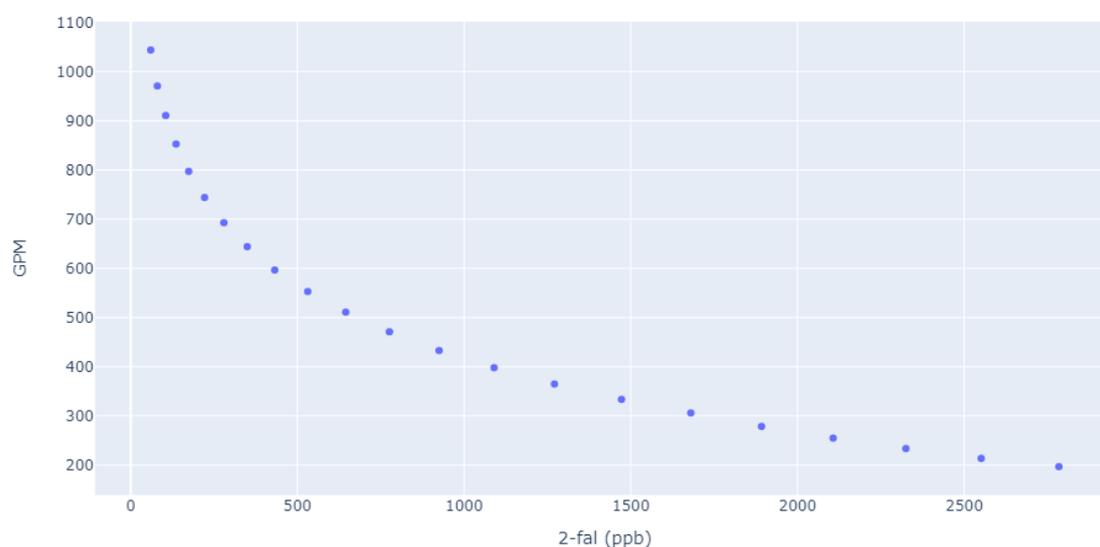


Figura 7: Relação entre GP e 2-fal

Como pode ser notado, a relação entre os dados aparenta se ajustar a uma curva do tipo exponencial decrescente. De fato, os dados parecem terem sido tratados e não serem os dados brutos das medições, pois não existem pontos que se deslocam para fora da curva. Podemos ainda aprofundar a análise entre esses dois atributos medindo a correlação

entre eles. Como os dados não se ajustam a uma reta, porém possuem a mesma tendência ao longo dos pontos, faz mais sentido, aqui, utilizar a correlação de Spearman. A Figura 8 apresenta essa análise.

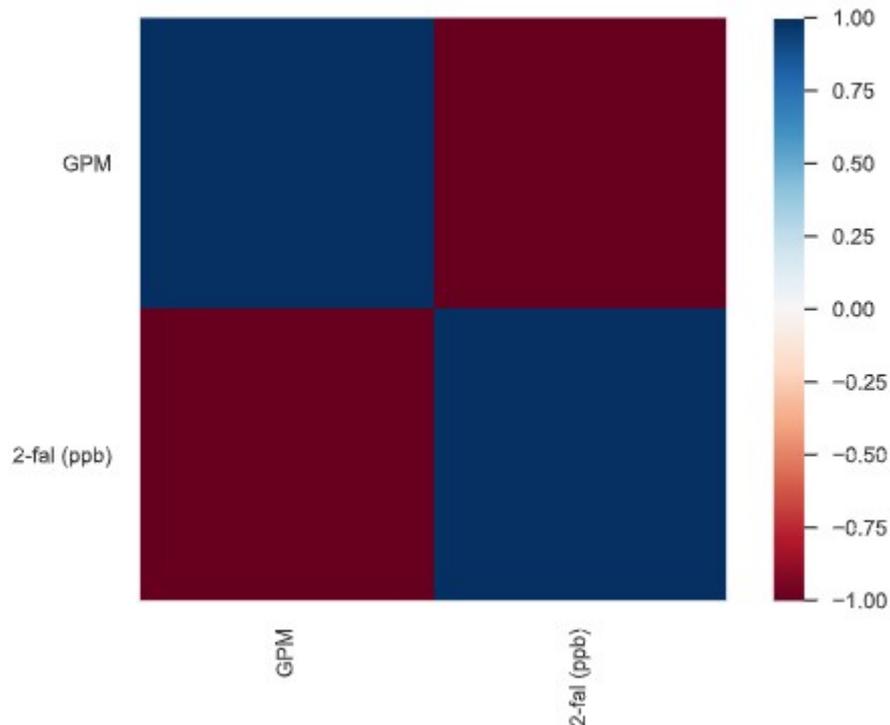


Figura 8: Matriz de correlação de Spearman

A correlação entre GP e 2-fal é fortemente negativa, indicando que o aumento no nível de 2-fal indica uma diminuição do GP. Conforme a revisão da literatura apresentada no capítulo anterior, esse comportamento era esperado, todavia não era esperada uma correlação tão forte (-1 para a correlação de Spearman e -0,9 para a correlação de Pearson).

### 3.1.3.2 Dados de (TEYMOURI; VAHIDI, 2017)

A descrição dos dados (TEYMOURI; VAHIDI, 2017) pode ser visualizada na Tabela 4.

Tabela 4: Descrição dos dados

	Potência (MVA)	CO <sub>2</sub> /CO	2-fal (ppb)	GPM
count	32,000000	32,000000	32,000000	32,000000
mean	86,531250	7,655312	413,093750	554,781250
std	39,240355	0,797840	420,861126	272,054161
min	10,000000	6,200000	10,000000	191,000000
50 %	95,500000	7,760000	262,000000	509,500000
max	154,000000	9,200000	1285,000000	1089,000000

Dados de 32 transformadores foram utilizados, sendo que suas potências variam de 10 MVA até 154 MVA, ou seja, transformadores de médio e grande porte foram utilizados. Além disso, a distribuição de GP para esse caso estava menos deslocada, uma vez que a média e a mediana estavam mais próximas. É possível verificar isso através do histograma apresentado na [Figura 9](#).

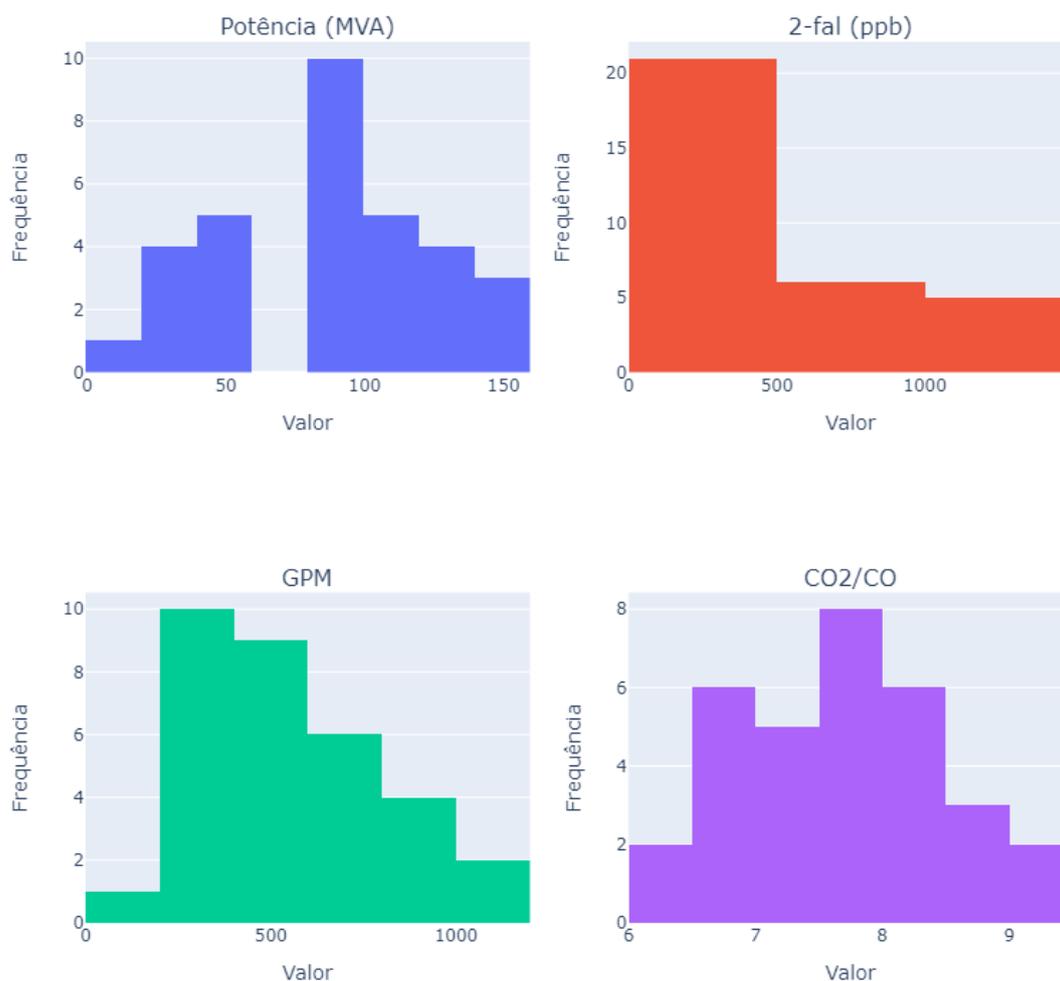


Figura 9: Histograma dos atributos

A relação entre 2-fal e GP, desse conjunto de dados, pode ser verificada na [Figura 10](#).

A relação entre esses dois atributos também se assemelha a uma exponencial negativa, porém os dados desse conjunto parecem ter sido menos tratados em comparação ao anterior, uma vez que o ajuste a uma curva não é perfeito.

Além disso, é possível também verificar a relação entre GP e a razão  $\text{CO}_2/\text{CO}$ , conforme a [Figura 11](#)

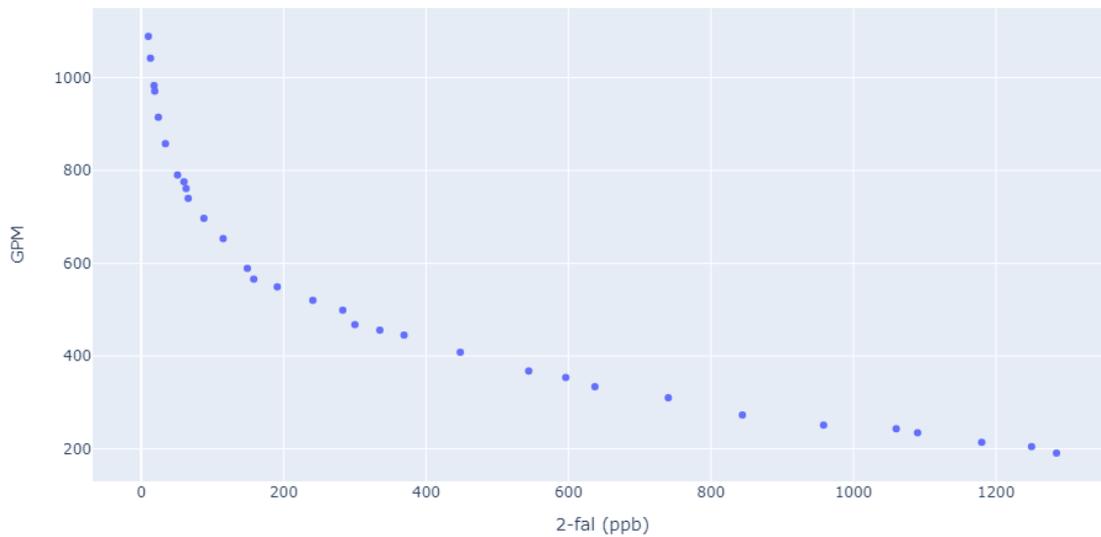
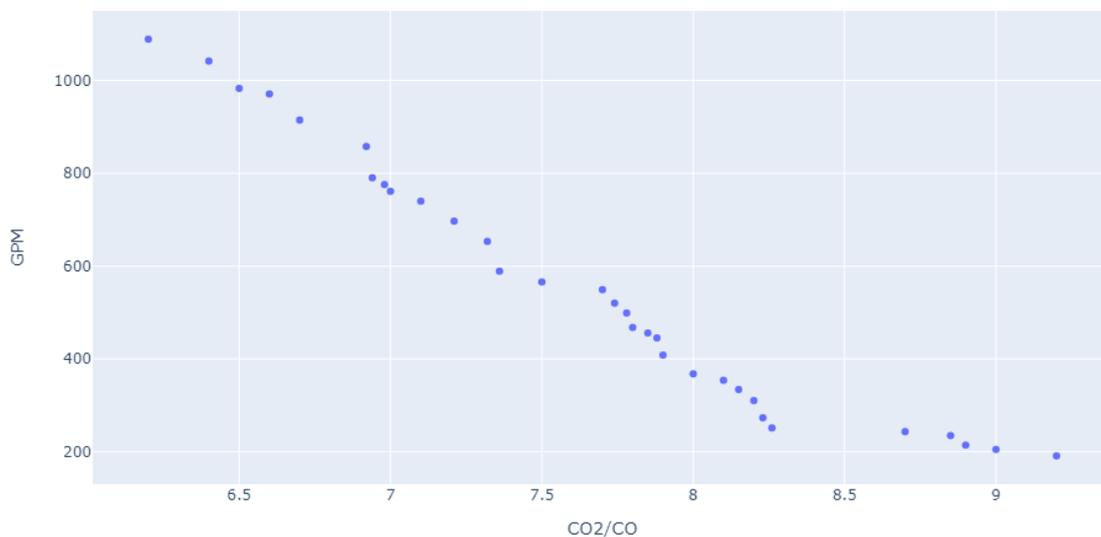


Figura 10: Relação GP e 2-fal

Figura 11: Relação GP e CO<sub>2</sub>/CO

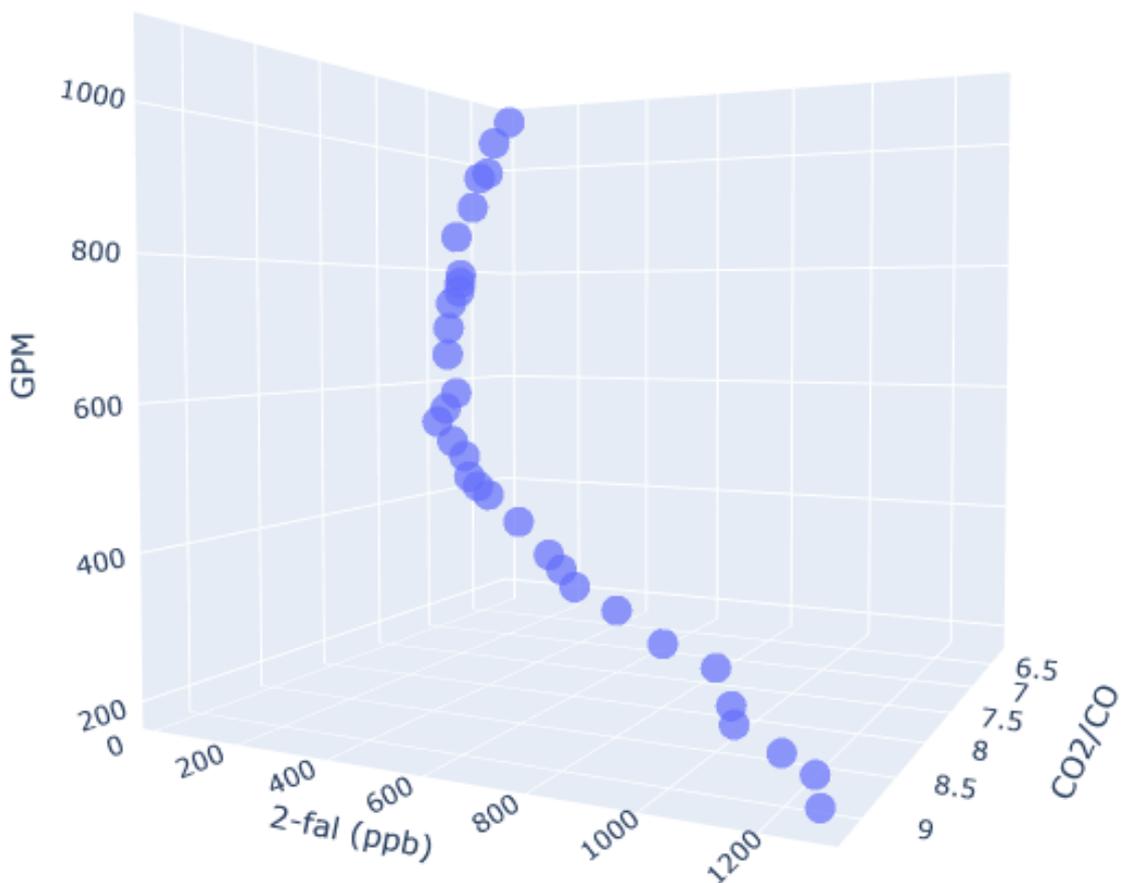
Isso é um forte indicativo de que o valor da razão CO<sub>2</sub>/CO também pode ser utilizada para prever o GP do transformador. Para analisarmos melhor essas relações, convém que uma matriz de correlação seja montada, conforme a [Tabela 5](#). Nela, é possível observar que a correlação entre CO<sub>2</sub>/CO e 2-fal possui um alto valor (0,93) e ambos apresentam alta correlação com o valor de GP. A alta correlação entre CO<sub>2</sub>/CO e 2-

fal é um indicativo de que ambos os atributos possuem informação redundante quando relacionados ao GP. Essa alta correlação está ligada ao fato de que ambos  $\text{CO}_2/\text{CO}$  e 2-fal são produtos diretos do envelhecimento do papel, sendo produzidos de forma semelhante por sua degradação (SAHA; PURKAIT, 2017). Todavia, como essa correlação não é perfeita, existe informação na razão  $\text{CO}_2/\text{CO}$  que não está presente na medição de 2-fal, tornando possível a utilização do  $\text{CO}_2/\text{CO}$  para realização de um ajuste fino em um modelo de aprendizagem de máquina.

Tabela 5: Matriz de Correlação

	Potência (MVA)	$\text{CO}_2/\text{CO}$	2-fal (ppb)	GPM
Potência (MVA)	1,000000	0,344027	0,446999	-0,264492
$\text{CO}_2/\text{CO}$	0,344027	1,000000	0,934144	-0,975286
2-fal (ppb)	0,446999	0,934144	1,000000	-0,879303
GPM	-0,264492	-0,975286	-0,879303	1,000000

A relação entre os principais atributos (GP, 2-fal e  $\text{CO}_2/\text{CO}$ ) pode ser visualizada na Figura 12.

Figura 12: Relação entre 2-fal,  $\text{CO}_2/\text{CO}$  e GP

#### 3.1.4 Conclusões da análise exploratória

Ambos os dados utilizados apresentam relação entre GP e 2-fal semelhante, próxima a uma exponencial decrescente. Os dados de (ASSUNÇÃO, 2007) possuem um comportamento mais filtrado enquanto os dados (TEYMOURI; VAHIDI, 2017) possuem um comportamento menos filtrado. Além disso, foi verificado que a razão  $\text{CO}_2/\text{CO}$ , por mais que possua alta correlação com o 2-fal, pode ser utilizada como atributo na previsão do grau de polimerização do transformador.

## 4 RESULTADOS OBTIDOS

### 4.1 Modelo de previsão do GP a partir do 2-fal

O primeiro modelo criado utilizou apenas o valor de 2-fal para prever o valor do GP. Para isso, foram utilizados os dados de [Assunção \(2007\)](#) e o problema foi tratado como um de regressão, no qual a saída GP foi calculada/predita a partir da entrada 2-fal. Após isso, o modelo foi testado com os dados de [Teymouri e Vahidi \(2017\)](#).

Para a realização do curvefitting, foi utilizada a função `polyfit`, da biblioteca `numpy`. Os dados foram divididos entre treinamento e validação, e o grau do polinômio foi definido através de um problema de otimização, cujo objetivo foi obter a maior média de erro médio ao longo dos folds. A proporção para treinamento e validação utilizada foi de 80 % para treinamento e 20 % para validação, pois como a quantidade de dados é pequena, aumentar a quantidade de dados para treinamento pode incorrer em um aumento de *overfitting*. Os erros médios para cada grau polinomial na validação podem ser visualizados na [Tabela 6](#).

Tabela 6: *Mean Average Error* ao longo dos *folds* para cada grau de polinômio

Grau	1	2	3	4	5	6	7	8	9	10	11	12	13
MAE	-25,10	-24,4	-24,7	-24,35	-23,22	-21,91	-20,66	-19,01	-18,76	-18,51	-14,33	-17,33	-33,69

Conforme é possível perceber, o menor erro médio foi obtido para o polinômio de grau 11.

Aplicando esse polinômio nos dados de teste, obteve-se um erro médio no valor de GP de 211. O resultado do treinamento e do teste podem ser visualizados nas [Figura 13](#) e [Figura 14](#).

Nos dados de teste percebe-se que existe um offset entre as curvas. Por mais que o formato seja bastante semelhante, esse offset leva a um erro médio de 211. Esse erro pode ser crítico principalmente para transformadores no fim de vida, uma vez que o valor predito indicaria um transformador em estado de deterioração avançado, mas não crítico. Dessa forma, faz-se necessária a consideração de mais atributos além do 2-fal para que a realização dessas previsões possa ser mais precisa.

### 4.2 Modelo de previsão do GP a partir do 2-fal e da razão CO<sub>2</sub>/CO

Como apresentado em ([TEYMOURI; VAHIDI, 2017](#)), a utilização da razão CO<sub>2</sub>/CO medido através de cromatografia pode ser utilizado para prever o GP de um transformador e possui alta taxa de correlação com o GP. Além disso, utilizar as medições de CO<sub>2</sub>/CO apresenta vantagens como:

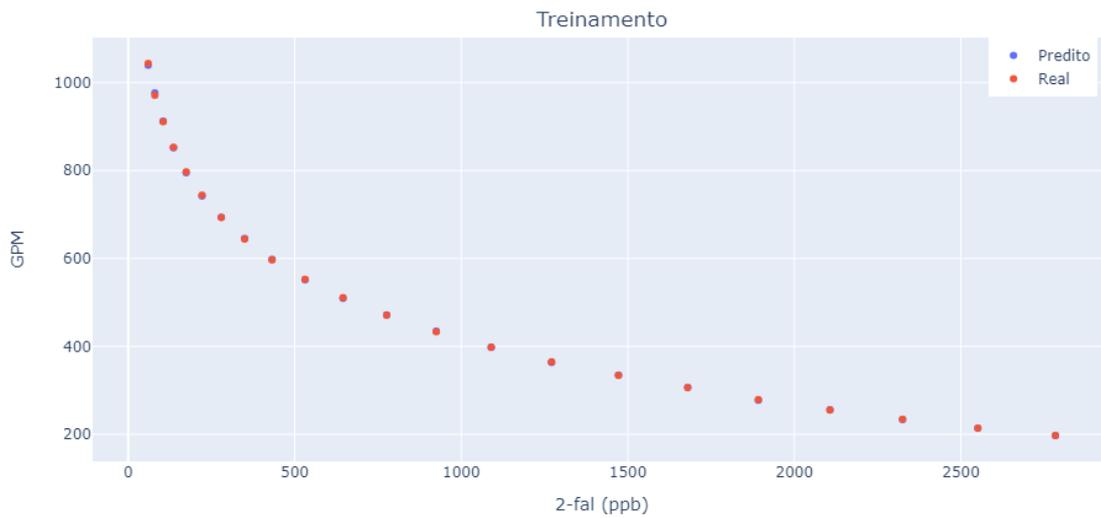


Figura 13: Relação entre GP predito e GP medido para o treinamento

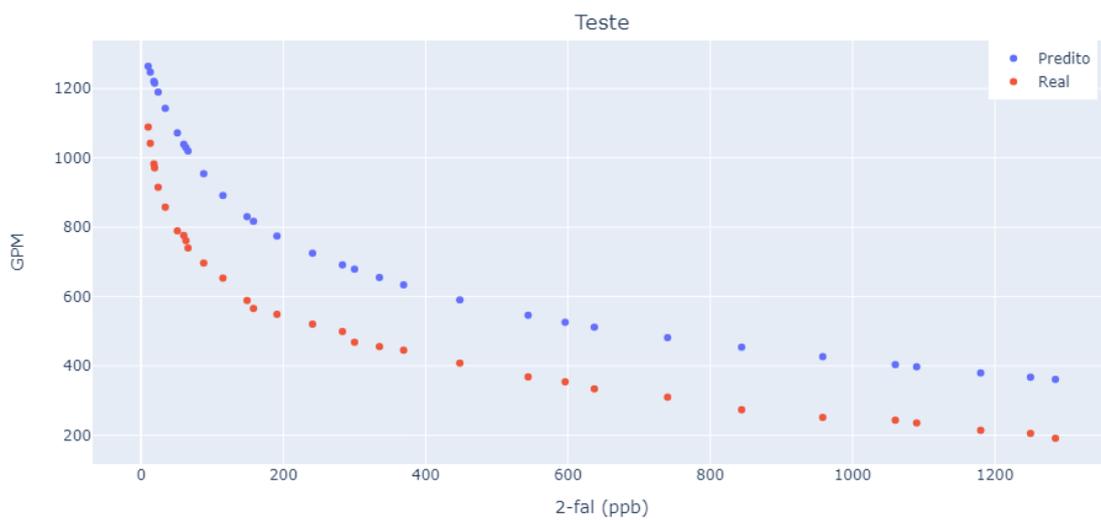


Figura 14: Relação entre GP predito e GP medido para o teste

- A proporção da relação  $\text{CO}_2/\text{CO}$  é mais estável;
- Se o vazamento de gás da atmosfera no transformador ocorre, a proporção não irá variar tanto quanto as concentrações individuais de  $\text{CO}_2$  e  $\text{CO}$ ;
- 2-fal é gerado em diferentes formas e, portanto, não é exclusivamente o resultado da degradação de celulose ([ALLAN, 1993](#));
- 2-fal não é estável em temperaturas acima  $110\text{ }^\circ\text{C}$ ;
- A presença de oxigênio, água e ácidos afeta a concentração de 2-fal ([PREVOST, 2005](#)).

Devido a isso, para esse segundo modelo, foi também utilizada a razão  $CO_2/CO$ . Diferentemente do primeiro modelo, no qual temos apenas um atributo de entrada (2-fal), para esse segundo modelo, convém aplicar *machine learning*. Como a quantidade de dados é pequena (apenas 32 transformadores), foi utilizado um modelo de aprendizagem de máquina simples. Além disso, devido à alta colinearidade (correlação) entre a relação  $CO_2/CO$  e 2-fal, foi necessária a aplicação de um modelo que penalize atributos com alta correlação. Por essa razão, foi escolhido o modelo Ridge Regression. Esse modelo é um tipo de regressão linear que possui um parâmetro  $\alpha$  que penaliza coeficientes grandes, impedindo que determinados coeficientes relacionados a atributos específicos recebam importância muito maior que outros, aumentando sua precisão. O parâmetro  $\alpha$  foi obtido através de validação cruzada, considerando 5 folds. As métricas obtidas no treinamento e no teste podem ser visualizadas na [Tabela 7](#).

Tabela 7: Métricas de avaliação do modelo utilizando os atributos: 2-fal e a relação  $CO_2/CO$

	<b>R2</b>	<b>MAE</b>
Treino	0,93	58,41
Teste	0,87	77,82

Conforme é possível perceber, ocorreu um pequeno *overfitting* aos dados de treinamento. Todavia o resultado apresentado pelo teste foi bastante satisfatório, atingindo um erro médio de 77,82, bem menor quando comparado ao primeiro modelo. Os resultados gráficos podem ser visualizados nas [Figura 15](#) e [Figura 16](#).

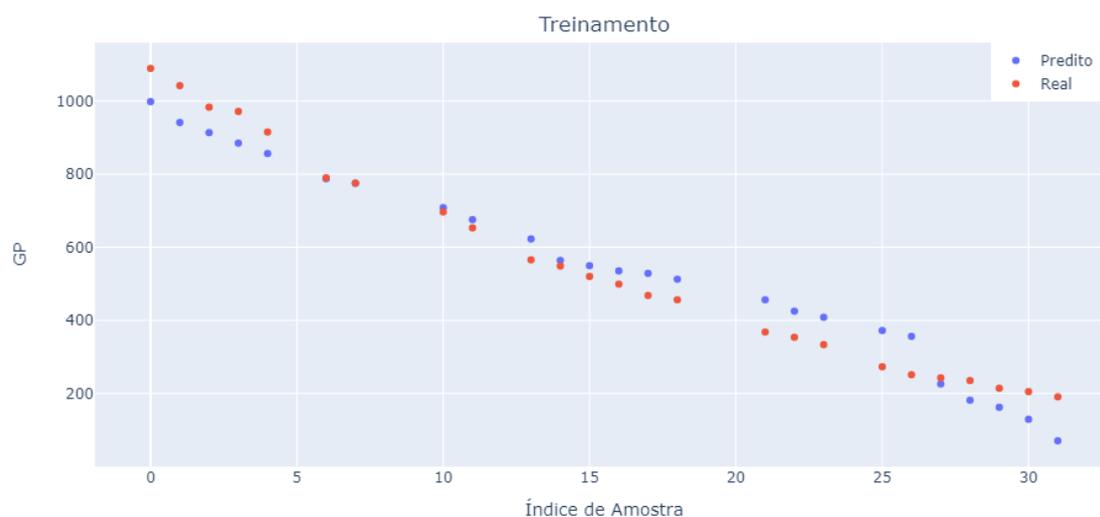


Figura 15: Relação entre GP predito e GP medido para o treinamento

Conforme é possível perceber o erro não é linear ao longo das amostras.

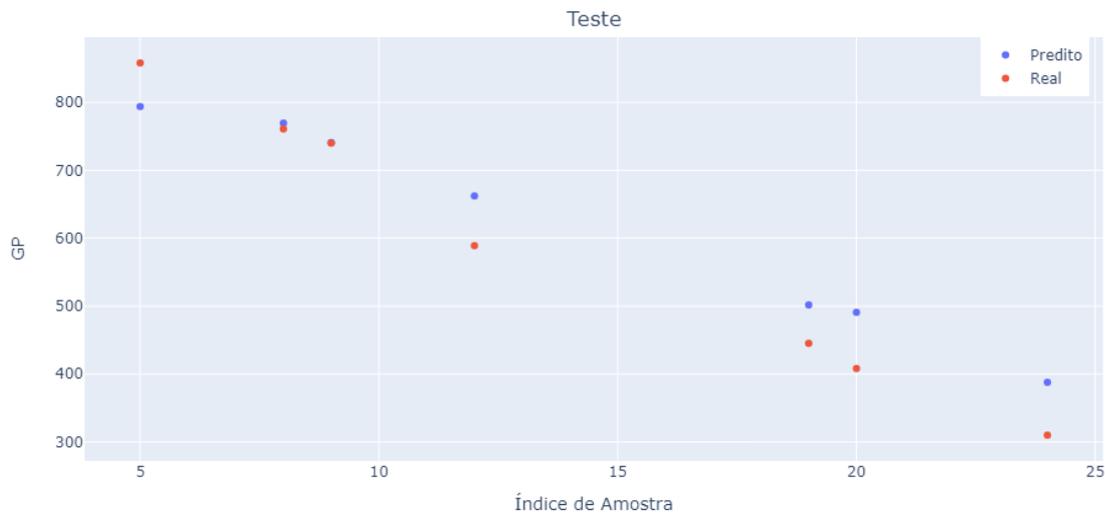


Figura 16: Relação entre GP predito e GP medido para o teste

### 4.3 Modelo de previsão do GP a partir do 2-fal, da razão $\text{CO}_2/\text{CO}$ e da fórmula de Chendong

Como a quantidade de dados utilizada para treinamento desse modelo é pequena, apenas a base de (TEYMOURI; VAHIDI, 2017), decidiu-se agregar conhecimento de outros artigos. Para isso, utilizou-se a fórmula de Chendong (Equação 4.1) como um atributo para o cálculo do GP real. A fórmula de Chendong possui uma razoável precisão para previsão do GP a partir das medições de 2-fal. Todavia, ainda assim, possui deficiências, principalmente por considerar apenas 2-fal.

$$GP = \frac{1,51 - \log(2fal)}{0,0035} \quad (4.1)$$

Desse modo, para esse modelo, os atributos de entrada considerados foram a medição de 2-fal em ppb, a medição da relação  $\text{CO}_2/\text{CO}$  e o GP calculado a partir da fórmula de Chendong. Os resultados são apresentados na Figura 17 na Figura 18.

Conforme é possível perceber, o resultado obtido foi significativamente melhor, estando bem mais próximo tanto para o treinamento quanto para o teste. A Figura 19 mostra os resultados apenas da equação de Chendong. O modelo desenvolvido obteve um resultado melhor que a equação de Chendong isolada. Esses resultados podem também ser verificados na Tabela 8.

O erro médio caiu significativamente do modelo 2 para o modelo 3, indicando que a adição do atributo Chendong agregou conhecimento útil ao modelo.

Tabela 8: Métricas de avaliação do modelo utilizando os atributos: 2-fal, relação  $CO_2/CO$  e GP calculado utilizando a fórmula de Chendong

	<b>R2</b>	<b>MAE</b>
Treinamento	0,99	5,84
Teste	0,99	8,37

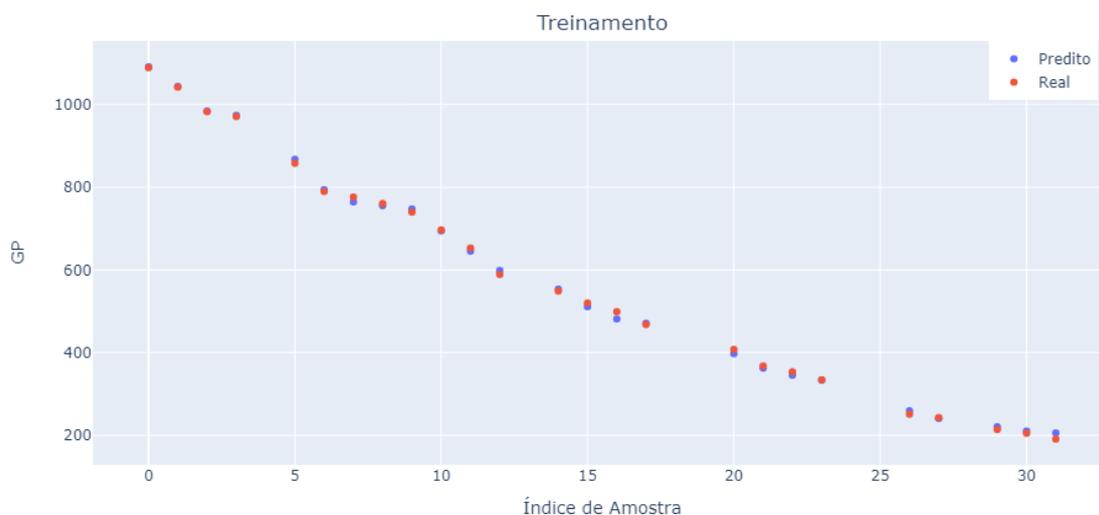


Figura 17: Relação entre GP predito e GP medido para o treinamento

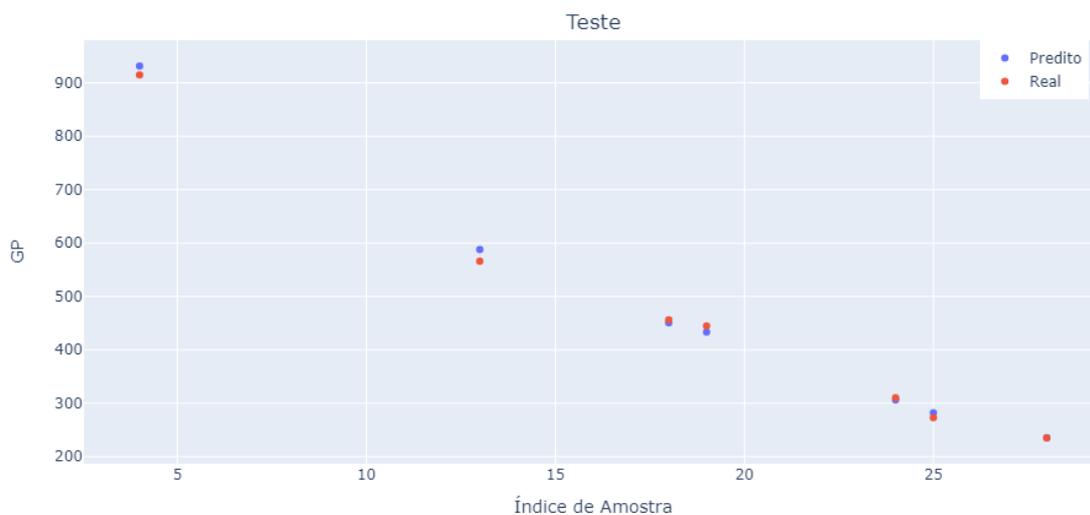


Figura 18: Relação entre GP predito e GP medido para o teste

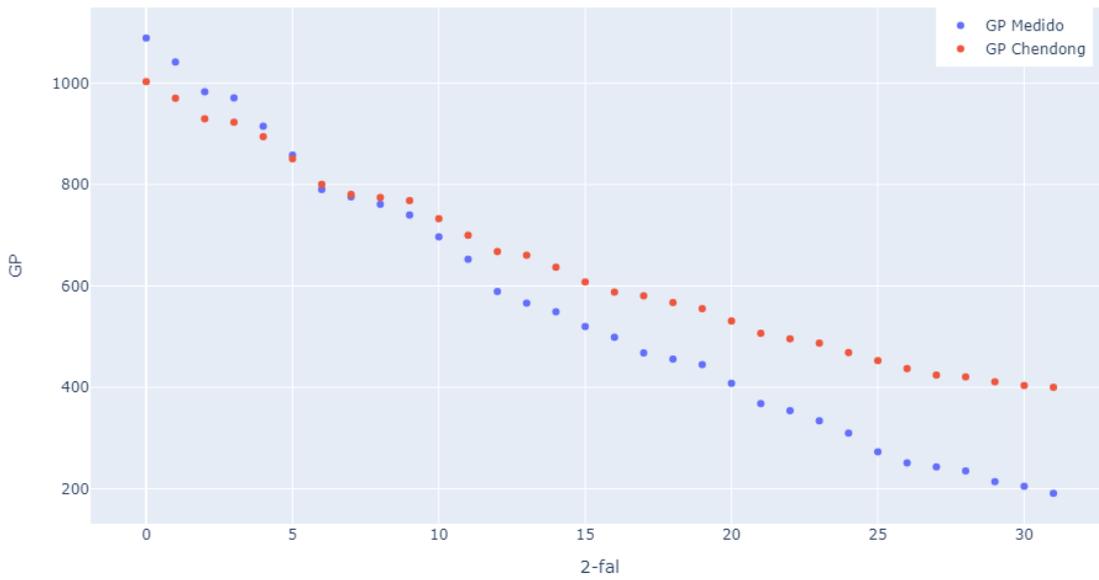


Figura 19: Relação entre GP real e GP Chendong

#### 4.4 Modelo de previsão do GP a partir do 2-fal, da relação $CO_2/CO$ , da fórmula de Chendong e da potência do transformador

Por fim, o último atributo adicionado ao modelo foi a potência do transformador. Por mais que esta se mostrasse sem correlação com o GP, os resultados obtidos mostraram que ela contém informação relevante para a previsão. Os resultados são apresentados na Figura 20 na Figura 21.

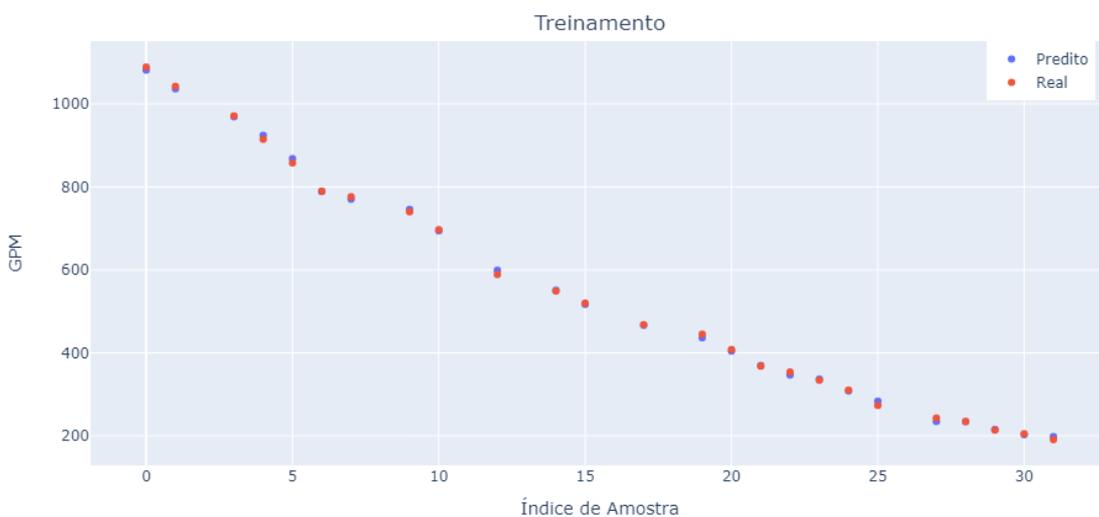


Figura 20: Relação entre GP predito e GP medido para o treinamento

A Tabela 9 apresenta esses resultados.

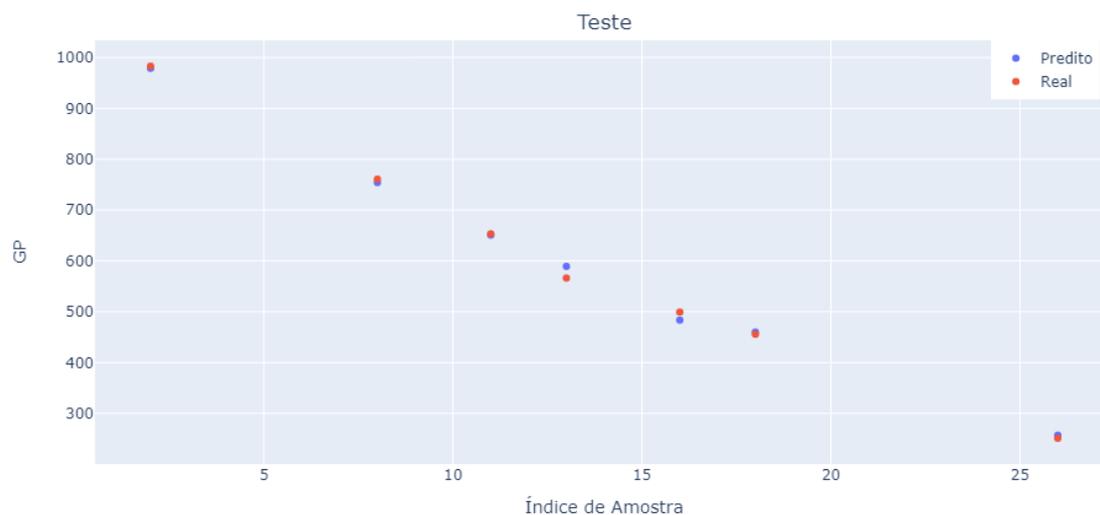


Figura 21: Relação entre GP predito e GP medido para o teste

Tabela 9: Métricas de avaliação do modelo utilizando os atributos: 2-fal, relação  $\text{CO}_2/\text{CO}$ , GP calculado utilizando a fórmula de Chendong e a Potência do Transformador

	<b>R2</b>	<b>MAE</b>
Treinamento	0,99	5,88
Teste	0,99	6,01

Conforme é possível perceber, o erro médio caiu ainda mais no teste. Todavia, como o último valor já estava suficientemente baixo, a diferença é quase imperceptível a nível de gráfico.

#### 4.5 Comparação dos modelos

Na tabela [Tabela 10](#) verifica-se os erros médios, em valor de GP, para cada modelo utilizados, tanto no treinamento, como no teste. Na [Figura 22](#) verifica-se que o aprimoramento do modelo, através da adição de outras variáveis, faz com que o erro médio caia significativamente.

Tabela 10: Comparação entre os modelos utilizados

<b>Modelo</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
MAE Treinamento	1,47	58,41	5,84	5,88
MAE Teste	211	77,82	8,37	6,01

O último modelo proposto mostrou ser o melhor modelo para previsão de vida útil do transformador, pois diminuiu ainda mais o erro médio na predição do valor do GP.

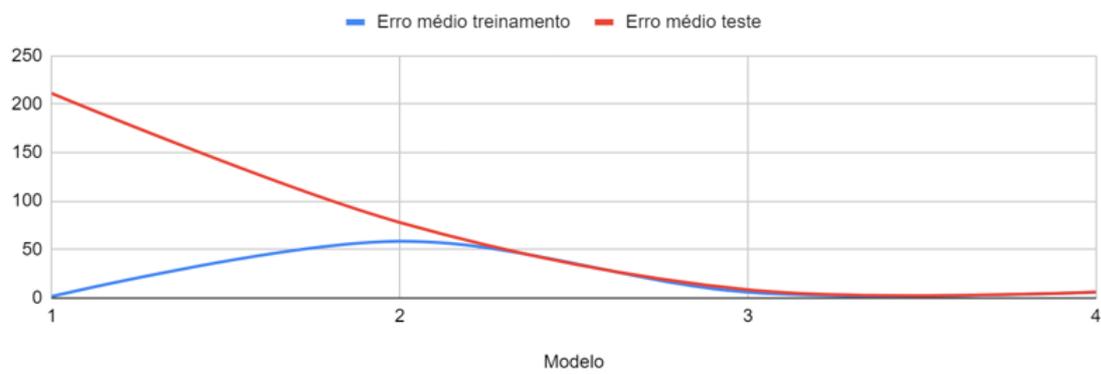


Figura 22: Erro modelos

## 5 CONCLUSÃO

Em decorrência do envelhecimento da isolação sólida do transformador, são gerados diversos possíveis indicadores de condição, entre eles, o 2-fal e  $\text{CO}_2/\text{CO}$  são os mais notáveis. O poder preditivo de cada um dos indicadores isoladamente não é suficiente para a criação de um modelo preciso que se ajuste aos diversos equipamentos existentes no mercado. Porém, quando tratados de forma unificada, esses dois indicadores compõem um modelo robusto com capacidade de prever os valores de GP de modo mais preciso.

Modelos de aprendizagem de máquina necessitam de uma grande quantidade de dados para funcionarem adequadamente, e por isso, foi também adotada a equação de Chendong como entrada para o modelo, contornando a falta de dados e possibilitando uma precisão ainda maior. Por fim, percebeu-se que a informação de potência do Transformador, por mais que não possua alta correlação com o valor de GP, apresenta conhecimento importante para o cálculo do envelhecimento utilizando o GP, auxiliando na diminuição do erro médio absoluto.

Devido à não disponibilidade de dados públicos desses indicadores, a continuidade destes estudos utilizando uma maior gama de dados está prevista. Primeiramente para que seja avaliada a importância de cada atributo na composição do valor final dos modelos, e posteriormente para que se possa utilizar uma maior taxa de treinamento dos modelos sem incorrer em um aumento de *overfitting*.

Apesar do modelo proposto não ter sido validado com um caso real distinto aos dados de teste, ele é expansível, e pode ser retreinado assim que novos dados estejam disponíveis. O modelo atual serve de ferramenta para auxílio a especialistas para uma melhor estimativa da vida útil dos transformadores utilizando o cálculo de GP.



## REFERÊNCIAS

- ACIU, A.-M. et al. Complementary analysis for dga based on duval methods and furan compounds using artificial neural networks. **Energies**, MDPI, v. 14, n. 3, p. 588, 2021.
- ALLAN, D. Practical life-assessment technique for aged transformer insulation. **IEE Proceedings A (Science, Measurement and Technology)**, IET, v. 140, n. 5, p. 404–408, 1993.
- ARVIND, D.; KHUSHDEEP, S.; DEEPAK, K. Condition monitoring of power transformer: A review. In: **2008 IEEE/PES Transmission and Distribution Conference and Exposition**. [S.l.: s.n.], 2008. p. 1–6.
- ASSUNÇÃO, T. C. B. N. **Contribuição à modelagem e análise do envelhecimento de transformadores de potência**. 2007. Tese (Doutorado) — Universidade Federal de Minas Gerais, 2007.
- CHALAPATHY, R.; CHAWLA, S. **Deep Learning for Anomaly Detection: A Survey**. 2019.
- DUVAL, M. et al. Significance and detection of very low degree of polymerization of paper in transformers. **IEEE Electrical Insulation Magazine**, v. 33, n. 1, p. 31–38, 2017.
- ELELE, U. et al. Towards online ageing detection in transformer oil: A review. **Sensors**, MDPI, v. 22, n. 20, p. 7923, 2022.
- IQBAL, M.; YAN, Z. Supervised machine learning approaches: A survey. **International Journal of Soft Computing**, v. 5, p. 946–952, 04 2015.
- KORDESTANI, M. et al. Failure prognosis and applications—a survey of recent literature. **IEEE Transactions on Reliability**, v. 70, n. 2, p. 728–748, 2021.
- LEAUPRASERT, K. et al. Intelligent machine learning techniques for condition assessment of power transformers. In: **2020 International Conference on Power, Energy and Innovations (ICPEI)**. [S.l.: s.n.], 2020. p. 65–68.
- LIN, Y. et al. Effect of oil replacement on furfural analysis and aging assessment of power transformers. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 22, n. 5, p. 2611–2619, 2015.
- MANDLIK, M.; RAMU, T. Moisture aided degradation of oil impregnated paper insulation in power transformers. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 21, n. 1, p. 186–193, 2014.
- MANGALATHU, S.; HWANG, S.-H.; JEON, J.-S. Failure mode and effects analysis of rc members based on machine-learning-based shapley additive explanations (shap) approach. **Engineering Structures**, v. 219, p. 110927, 2020. ISSN 0141-0296. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0141029620307513>>.
- MCDONALD, G. C. Ridge regression. **WIREs Computational Statistics**, v. 1, n. 1, p. 93–100, 2009. Disponível em: <<https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wics.14>>.

- MENTLIK, V. et al. On-line diagnostics of power transformers. In: **Conference Record of the 2006 IEEE International Symposium on Electrical Insulation**. [S.l.: s.n.], 2006. p. 546–549.
- MUHAMMAD, I.; YAN, Z. Supervised machine learning approaches: A survey. In: **SOCO 2015**. [S.l.: s.n.], 2015.
- MÜNSTER, T. et al. Investigation on ageing parameters of a thermally accelerated aged paper-oil-insulation in a hermetically sealed system. In: **2018 Condition Monitoring and Diagnosis (CMD)**. [S.l.: s.n.], 2018. p. 1–6.
- OMAR, S. et al. Machine learning techniques for anomaly detection: An overview. **International Journal of Computer Applications**, v. 79, 10 2013.
- ORIA, C. et al. State-of-the-art review on the performance of cellulosic dielectric materials in power transformers: Mechanical response and ageing. **IEEE Transactions on Dielectrics and Electrical Insulation**, IEEE, v. 26, n. 3, p. 939–954, 2019.
- PREVOST, T. A. Thermally upgraded insulation in transformers. In: IEEE. **Proceedings Electrical Insulation Conference and Electrical Manufacturing Expo, 2005**. [S.l.], 2005. p. 120–125.
- SAHA, T. K.; PURKAIT, P. **Transformer ageing: monitoring and estimation techniques**. [S.l.]: John Wiley & Sons, 2017.
- SHUTENKO, O.; PONOMARENKO, S. Analysis of ageing characteristics of transformer oils under long-term operation conditions. **Iranian Journal of Science and Technology, Transactions of Electrical Engineering**, Springer, v. 46, n. 2, p. 481–501, 2022.
- TEYMOURI, A.; VAHIDI, B. Co<sub>2</sub>/co concentration ratio: A complementary method for determining the degree of polymerization of power transformer paper insulation. **IEEE Electrical Insulation Magazine**, IEEE, v. 33, n. 1, p. 24–30, 2017.
- VILAKAZI, C.; MARWALA, T. Bushing fault detection and diagnosis using extension neural network. In: **2006 International Conference on Intelligent Engineering Systems**. [S.l.: s.n.], 2006. p. 170–174.
- ZHANG, D. et al. A bi-level machine learning method for fault diagnosis of oil-immersed transformers with feature explainability. **International Journal of Electrical Power Energy Systems**, v. 134, p. 107356, 2022. ISSN 0142-0615. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0142061521005950>.