



**UNIVERSIDADE FEDERAL DO PIAUÍ**  
**CENTRO DE TECNOLOGIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**  
**MESTRADO ACADÊMICO EM ENGENHARIA ELÉTRICA**

**ABRAÃO GALENO DA COSTA MENEZES**

**ÁRVORES DE DECISÃO PARA DIAGNÓSTICO DE FALHAS INCIPIENTES EM  
TRANSFORMADORES DE POTÊNCIA: NOVAS PERSPECTIVAS E REVISÃO DO  
TRIÂNGULO DE DUVAL**

**TERESINA**

**2018**

ABRAÃO GALENO DA COSTA MENEZES

ÁRVORES DE DECISÃO PARA DIAGNÓSTICO DE FALHAS INCIPIENTES EM  
TRANSFORMADORES DE POTÊNCIA: NOVAS PERSPECTIVAS E REVISÃO DO  
TRIÂNGULO DE DUVAL

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia Elétrica do Programa de Pós-Graduação em Engenharia Elétrica do Centro de Tecnologia da Universidade Federal do Piauí, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica. Área de Concentração: Sistemas de Energia

Orientador: Prof. Dr. Otacílio da Mota Almeida

Co-Orientador: Prof. Dr. Fábio Rocha Barbosa

TERESINA

2018

**ABRAÃO GALENO DA COSTA MENEZES**

**ÁRVORES DE DECISÃO PARA DIAGNÓSTICO DE FALHAS INCIPIENTES EM  
TRANSFORMADORES DE POTÊNCIA: NOVAS PERSPECTIVAS E REVISÃO DO  
TRIÂNGULO DE DUVAL**

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia Elétrica, Área de Concentração em Sistemas de Energia Elétrica, Linha de Pesquisa Automação e Controle, e aprovada em sua forma final pelo Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Piauí.

---

Abraão Galeno da Costa Menezes

**Orientador:**

---

Otacílio da Mota Almeida, Dr.

**Banca Examinadora:**

---

Prof. Luis Gustavo Mota Souza, Dr.

---

Prof. Hermes Manoel Galvão Castelo Branco, Dr.

---

Prof. Osvaldo Ronald Saavedra Mendez, Dr.

Teresina, 23 de agosto de 2018

Aos meus filhos Abraão e Rafael, que pequenos ainda, não sabem a importância de seus sorrisos, carinhos e abraços para a minha caminhada.

## AGRADECIMENTOS

A Deus, por proporcionar a dádiva de acordar dia após dia, guiando meus caminhos e norteando meus pensamentos.

Aos meus filhos Abraão Galeno da Costa Menezes Filho e Rafael Poluca Galeno Menezes, razões da minha existência e persistência.

À minha esposa, Rafaela Poluca Galeno Menezes, por me acompanhar nesta caminhada.

Aos meus pais Raimundo Araújo Menezes e Maria Cecília Galeno da Costa, por serem meus exemplos, alvos a serem alcançados.

Aos meus familiares que presentes ou não, fizeram parte desta caminhada.

Ao Prof. Dr. Otacílio da Mota Almeida por me orientar na minha dissertação de mestrado.

Ao Prof. Dr. Fábio da Rocha Barbosa, co-orientador da minha dissertação.

Ao Prof. Dr. Antônio Airton Carneiro de Freitas, por me orientar no caminho da pesquisa e incentivo pessoal.

Ao Prof. Dr. Aryfrance Rocha Almeida, pelas explanações técnicas e troca de informações.

A todos os professores do Programa de Pós-graduação em Engenharia Elétrica-PPGEE da Universidade Federal do Piauí que juntos tornaram-se presença indispensável na minha formação.

Aos meus colegas de turma, pois juntos dividimos conhecimentos, angústias, descobertas e principalmente amizades.

Aos meus alunos, pela curiosidade, perguntas desconcertantes e acima de tudo vontade de saber.

A Universidade Federal do Piauí-UFPI, ao Centro de Tecnologia-CT e ao Programa de Pós-Graduação em Engenharia Elétrica-PPGEE representado pelo Prof. Dr. Otacílio da Mota Almeida pela oportunidade da melhoria.

“Aprendi que a coragem não é a ausência do medo, mas o triunfo sobre ele. O homem corajoso não é aquele que não sente medo, mas o que conquista esse medo.”

(Nelson Mandela)

## RESUMO

O diagnóstico de falhas incipientes em transformadores de potência, com base na análise dos gases dissolvidos no óleo isolante, é o método mais utilizado em ocasiões que envolvam manutenções preventivas, preditivas ou corretivas no equipamento. É essencial garantir a operação contínua do transformador de força e evitar possíveis falhas que possam ocorrer devido ao seu ciclo de vida natural ou arranjo elétrico a que são submetidos. Atualmente, o método do triângulo de Duval é uma das técnicas tradicionais mais utilizadas na Análise de Gases Dissolvidos (DGA, do Inglês *Dissolved Gas Analysis*), porém, esta técnica tem mostrado precisão limitada. Para superar os problemas de desempenho convencionais, técnicas de Inteligência Computacional como Redes Neurais, Sistemas Difusos e, mais recentemente, Árvores de Decisão (DT, do inglês *Decision Tree*) foram propostas como métodos para melhorar a análise de gases dissolvidos. Este trabalho mostra que o algoritmo de Árvores de Decisão, usando a taxa de ganho como uma métrica para seleção de atributos, foi capaz de extrair o máximo de informação destes, e fornecer uma solução para casos não resolvidos usando métodos tradicionais. Mostrou-se também que os atributos mapeados pela árvore de decisão para classificação DGA podem inferir no tipo de falha através da análise de poucos gases. Finalmente, pode-se concluir que a técnica proposta desempenha um papel importante na melhoria da análise DGA e configura-se como uma ferramenta promissora utilizada isolada, ou em conjunto com as técnicas tradicionais.

**Palavras-chave:** Transformador de Força. DGA. Árvores de Decisão.

## ABSTRACT

The diagnosis of incipient fault in power transformers, based on analysis of the dissolved gases in mineral oil, the most used method on occasions involving preventive, predictive or corrective maintenance on equipment. It is essential to ensure the power transformer continuous operation and prevent possible failures that may occur because of their natural life cycle or electrical arrangement that are submitted. Currently, Duval Triangle method is one of the most used traditional techniques for Dissolved Gas Analysis, however this technique has shown limited accuracy. To overcome the conventional performance problems, Computational Intelligence techniques as neural networks, fuzzy systems and more recently Decision Trees have been proposed as methods for DGA analysis. This work proves that Decision Tree algorithm, by using the gain rate as a metric for attribute selection, have been able to extract as much information as possible from each attribute, and provide a solution for unsolved cases using traditional diagnose method. Showed the attributes mapped by the decision tree for DGA classification may inferred in the type of fault by the analysis of few gases. Finally, it can be concluded that the Decision Trees technique plays an important role for improving DGA analysis and it appears to be a promising tool by itself or in conjunction with traditional techniques.

**Keywords:** Power transformer.DGA.Decision trees.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Destilação fracionada do petróleo . . . . .	25
Figura 2 – Cadeia Naftênica . . . . .	26
Figura 3 – Cadeia Parafínica Linear . . . . .	26
Figura 4 – Cadeia Parafínica Ramificada . . . . .	27
Figura 5 – Reações de Pirólise do OMI. . . . .	28
Figura 6 – Formação dos gases em função da temperatura de decomposição. a) pressão x temperatura. b) concentração dos gases x temperatura. . . . .	29
Figura 7 – Transformador de força 69/13.8 kV com a potência de 25 MVA avariado em ocorrência de falha elétrica . . . . .	32
Figura 8 – Avaria provocada por falha elétrica na parte inferior do enrolamento central do transformador . . . . .	32
Figura 9 – Diagnóstico: Falha térmica no óleo . . . . .	35
Figura 10 – Diagnóstico: Falha térmica na celulose . . . . .	36
Figura 11 – Diagnóstico: Descarga parcial ou Corona . . . . .	36
Figura 12 – Diagnóstico: Descarga de alta energia . . . . .	37
Figura 13 – Gráfico de Dornenburg . . . . .	39
Figura 14 – Triângulo de Duval . . . . .	41
Figura 15 – Diagramas fuzzy para as razões de Rogers . . . . .	46
Figura 16 – Sistema de inferência para diagnóstico de falhas em transformadores para Rogers . . . . .	47
Figura 17 – Descoberta do Conhecimento em Base de Dados segundo(FAYYAD et al., 1996b) . . . . .	51
Figura 18 – Taxonomia do método de Data Mining . . . . .	54
Figura 19 – Exemplo de Árvore de Decisão . . . . .	55
Figura 20 – Partição do espaço de variáveis e regras obtidas de uma árvore de classificação binária . . . . .	57
Figura 21 – Exemplo de pseudocódigo para construção de Árvore de Decisão . . . . .	59
Figura 22 – Estrutura de árvore para os objetos em C . . . . .	67
Figura 23 – Exemplo de pseudo-código para o C4.5 . . . . .	70
Figura 24 – Espaço amostral . . . . .	74
Figura 25 – Novo Espaço de entrada . . . . .	76

Figura 26 – 1° e 2° Nós da Árvore de Decisão . . . . .	80
Figura 27 – Árvore 1 . . . . .	82
Figura 28 – Erro 01 árvore 1 . . . . .	83
Figura 29 – Erro 02 Árvore 1 . . . . .	84
Figura 30 – Erro 03 Árvore 1 . . . . .	84
Figura 31 – Árvore 2 . . . . .	85
Figura 32 – Modificação ocorrida da árvore 01 para a 02 . . . . .	86
Figura 33 – Árvore Final . . . . .	88
Figura 34 – Modificação ocorrida da árvore 2 para a 3 . . . . .	89
Figura 35 – Rastreabilidade da porcentagem de acetileno ( $PC_2H_2$ ) . . . . .	93
Figura 36 – Divisão do acetileno segundo a teoria de formação dos gases dissolvidos . . . . .	94
Figura 37 – Rastreabilidade da porcentagem de etileno ( $PC_2H_4$ ) . . . . .	94
Figura 38 – Divisão do etileno segundo a teoria de formação dos gases dissolvidos . . . . .	95
Figura 39 – Rastreabilidade da porcentagem de metano ( $PCH_4$ ) . . . . .	96
Figura 40 – Atributo metano segundo a teoria de formação dos gases dissolvidos . . . . .	97
Figura 41 – Rastreabilidade do hidrogênio ( $H_2$ ) . . . . .	98
Figura 42 – Atributo hidrogênio segundo a teoria de formação dos gases dissolvidos . . . . .	99
Figura 43 – Rastreabilidade do metano ( $CH_4$ ) . . . . .	99
Figura 44 – Rastreabilidade do monóxido de carbono( $CO$ ) . . . . .	101
Figura 45 – Atributos monóxido e dióxido de carbono para amostras classificadas como falhas térmicas e descargas parciais . . . . .	102
Figura 46 – Rastreabilidade do dióxido de carbono ( $CO_2$ ) . . . . .	103
Figura 47 – Rastreabilidade do etano ( $C_2H_6$ ) . . . . .	104
Figura 48 – Atributo etano segundo a teoria de formação dos gases dissolvidos . . . . .	105
Figura 49 – Rastreabilidade da relação( $C_2H_4/CH_4$ ) . . . . .	105
Figura 50 – Rastreabilidade da relação( $CH_4/H_2$ ) . . . . .	106
Figura 51 – Relação ( $CH_4/H_2$ ) segundo a teoria de formação dos gases dissolvidos . . . . .	107
Figura 52 – Triângulo de Duval Clássico x Triângulo Proposto. . . . .	110
Figura 53 – $\%C_2H_2$ x $\%C_2H_4$ . . . . .	112
Figura 54 – Regra do Acetileno para os quatro casos em Siada e Hmood, (2013) . . . . .	114
Figura 55 – Ramos da DT proposta utilizados na classificação das amostras em Wang (2014) . . . . .	116

Figura 56 – Ramos da DT proposta Utilizados na classificação das amostras de falhas  
elétricas em (WANG,2014) . . . . . 118

## LISTA DE TABELAS

Tabela 1 – Características normais dos óleos parafínicos e naftênicos . . . . .	27
Tabela 2 – Relação entre gases para DGA . . . . .	38
Tabela 3 – Diagnostico pelo método de Dornenburg . . . . .	38
Tabela 4 – Limites mínimos para validação do método de Dornenburg . . . . .	38
Tabela 5 – Diagnostico pelo método de Rogers . . . . .	40
Tabela 6 – Etapas do Processo de Descoberta do Conhecimento . . . . .	52
Tabela 7 – Escolha do Atributo para o Nó Raiz da Árvore de Decisão . . . . .	78
Tabela 8 – Divisão das Amostras pelo Nó-raiz . . . . .	79
Tabela 9 – Escolha do Atributo para o 2º Nó Raiz da Árvore de Decisão . . . . .	79
Tabela 10 – Matriz de Confusão para classificação DGA utilizando a árvore 1 . . . . .	81
Tabela 11 – Matriz de Confusão para classificação DGA utilizando a árvore 2 . . . . .	85
Tabela 12 – Matriz de Confusão para classificação DGA utilizando a árvore final . . . . .	87
Tabela 13 – Novo espaço de entradas para classificação DGA . . . . .	89
Tabela 14 – Regras extraídas pela Árvore Proposta para Classificação DGA . . . . .	108
Tabela 15 – Percentual de Acerto na Classificação DGA Comparado a outros Métodos . . . . .	108
Tabela 16 – Características da DT Induzidas pelo Algorimto C4.5 utilizando os parâmetros do Triângulo de Duval. . . . .	109
Tabela 17 – Matriz de Confusão para Classificação DGA Utilizando a Árvore Elaborada com os Parâmetros do Triângulo de Duval . . . . .	109
Tabela 18 – Regras extraídas pela Árvore Proposta para Classificação DGA . . . . .	110
Tabela 19 – Comparação na classificação: Método Fuzzificado x DT Proposta . . . . .	113
Tabela 20 – Diagnóstico do Acompanhamento do transformador em falha térmica pela DT proposta . . . . .	115
Tabela 21 – Diagnóstico do acompanhamento do transformador em falha elétrica pela DT proposta . . . . .	117

## LISTA DE SÍMBOLOS

$\Theta$	Operação lógica testada(=, $\neq$ , $\leq$ , $\geq$ )
$\alpha$	Valor limite
$c_1$ e $c_2$	"Caminhos" distintos em uma árvore de decisão
$S$	Conjunto de treinamento para uma árvore de decisão
$S_t$	Subconjunto do conjunto de treinamento
$A$	Atributo do conjunto de entrada
$E$	Entropia
$G$	Ganho de Informação
$x$	Variável aleatória
$K$	Valor discreto
$\Phi$	Medida de impureza
$P_y$	Vetor probabilidade do atributo $y$
$ORT$	Critério Ortogonal para seleção de atributo
$T$	Sequência de árvores criadas
$\varepsilon$	Taxa de erro da árvore $T$
$gain(A)$	Ganho obtido pelo atributo "A"

## LISTA DE ABREVIATURAS E SIGLAS

ANEEL	Agencia Nacional de Energia Elétrica
IEEE	<i>Institute of Electrical and Electronic Engineers</i>
IEC	<i>International Electrotechnical Commission</i>
DGA	<i>Dissolved Gas Analysis</i> – Análise de gases dissolvidos
DT	<i>Decision Tree</i> – Árvore de Decisão
AC	Árvore de Classificação
ACB	Árvore de Classificação Binária
AR	Árvore de Regressão
IC	Inteligência Computacional
OMI	Óleo Mineral Isolante
CP	Conteúdo Parafínico
RNA	Rede Neurais Artificial
MLP	<i>Multilayer Perceptron</i>
ML	<i>Machine Learning</i> – Aprendizagem de Máquina
SIF	Sistema de Inferência <i>Fuzzy</i>
PD	Descargas Parciais
D1	Descargas de baixa energia
D2	Descargas de alta energia
T1	Falhas térmicas de temperatura $< 300^{\circ}\text{C}$
T2	Falhas térmicas de temperaturas $300^{\circ}\text{C} < T < 700^{\circ}\text{C}$
T3	Falhas térmicas de temperaturas $> 700^{\circ}\text{C}$
OLTC	<i>On Load Tap Changer</i>
ppm	Partes por milhão
ABNT	Associação Brasileira de Normas Técnicas
ASTM	<i>American Society for Testing and Materials</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	17
<b>1.1</b>	<b>Motivação para o desenvolvimento do tema abordado</b>	17
<i>1.1.1</i>	<i>O Contexto Atual</i>	17
<i>1.1.2</i>	<i>Motivação para a Pesquisa</i>	19
<b>1.2</b>	<b>Objetivos da Pesquisa</b>	20
<b>1.3</b>	<b>Disposição Geral do documento</b>	21
<b>1.4</b>	<b>Produção Científica</b>	22
<b>2</b>	<b>FALHAS INCIPIENTES EM TRANSFORMADORES DE FORÇA</b>	23
<b>2.1</b>	<b>O óleo isolante em transformadores de força</b>	23
<i>2.1.1</i>	<i>Óleo mineral isolante(OMI)</i>	24
<i>2.1.1.1</i>	<i>Óleo mineral isolante naftênico</i>	25
<i>2.1.1.2</i>	<i>Óleo mineral isolante parafínico</i>	25
<i>2.1.2</i>	<i>Comparação entre os óleos parafínicos e naftênicos</i>	26
<i>2.1.3</i>	<i>Degradação dos óleos minerais isolantes</i>	26
<i>2.1.3.1</i>	<i>Processo de oxidação dos óleos minerais isolantes</i>	27
<i>2.1.3.2</i>	<i>Degradação Térmica dos Óleos Minerais</i>	28
<b>2.2</b>	<b>Falhas em transformadores</b>	29
<i>2.2.1</i>	<i>Falhas Elétricas</i>	30
<i>2.2.2</i>	<i>Falhas Térmicas</i>	31
<b>2.3</b>	<b>Métodos convencionais utilizados para análise de gases dissolvidos no óleo do transformador de força</b>	34
<i>2.3.1</i>	<i>Análise do Gás Chave</i>	34
<i>2.3.1.1</i>	<i>Falha térmica no óleo</i>	34
<i>2.3.1.2</i>	<i>Falha térmica na celulose</i>	35
<i>2.3.1.3</i>	<i>Descarga parcial ou Corona</i>	35
<i>2.3.1.4</i>	<i>Descarga de Alta Energia ou Arco</i>	35
<i>2.3.2</i>	<i>Razões de Dornenburg</i>	37
<i>2.3.3</i>	<i>Razões de Rogers</i>	39
<i>2.3.4</i>	<i>Triângulo de Duval</i>	40

2.4	<b>Limitações impostas pelos métodos de análise de gases dissolvidos convencionais</b> . . . . .	41
2.5	<b>Propostas atuais para análise de gases dissolvidos no óleo do transformador de força baseados em Inteligência Computacional</b> . . . . .	43
2.5.1	<i>Redes Neurais</i> . . . . .	44
2.5.2	<i>Sistemas de inferência Fuzzy</i> . . . . .	45
2.5.3	<i>Árvores de decisão</i> . . . . .	47
2.6	<b>Considerações Parciais</b> . . . . .	49
3	<b>DATA MINING E ÁRVORES DE DECISÃO</b> . . . . .	50
3.1	<b>Descoberta de Conhecimento em Base de Dados</b> . . . . .	50
3.1.1	<i>Data Mining</i> . . . . .	51
3.2	<b>Árvore de Decisão</b> . . . . .	54
3.2.1	<i>Indução de Árvores de Decisão</i> . . . . .	57
3.2.2	<i>Critérios utilizados para divisão de Árvores de Decisão</i> . . . . .	60
3.2.2.1	<i>Critérios baseados em impurezas</i> . . . . .	60
3.2.2.2	<i>Ganho de Informação</i> . . . . .	61
3.2.2.3	<i>Índice Gini</i> . . . . .	61
3.2.2.4	<i>Razão do Ganho</i> . . . . .	62
3.2.2.5	<i>Critério Ortogonal</i> . . . . .	62
3.2.3	<i>Poda em Árvores de Decisão</i> . . . . .	63
3.2.3.1	<i>Poda por Custo de Complexidade</i> . . . . .	64
3.2.3.2	<i>Erro mínimo de Poda</i> . . . . .	64
3.3	<b>Algoritmos para Indução de Árvores de Decisão</b> . . . . .	65
3.3.1	<i>Algoritmo ID3</i> . . . . .	65
3.3.1.1	<i>Indução da árvore no ID3</i> . . . . .	66
3.3.2	<i>Algoritmo C4.5</i> . . . . .	68
3.3.2.1	<i>Indução de árvores no Algoritmo C4.5</i> . . . . .	69
3.4	<b>Considerações Parciais</b> . . . . .	71
4	<b>UTILIZAÇÃO DE ÁRVORES DE DECISÃO PARA CLASSIFICAÇÃO</b>	
	<b>DGA</b> . . . . .	73
4.1	<b>Seleção e tratamento de dados</b> . . . . .	73
4.1.1	<i>Aumento da dimensionalidade do espaço de entrada</i> . . . . .	74

4.1.2	<i>Pré-processamento e limpeza dos dados</i> . . . . .	77
4.2	<b>Árvores de decisão para classificação DGA</b> . . . . .	80
4.2.1	<i>Árvore 01</i> . . . . .	80
4.2.2	<i>Árvore 02</i> . . . . .	83
4.2.3	<i>Árvores 03, 04 e 05</i> . . . . .	87
4.3	<b>Composição do novo espaço de entradas</b> . . . . .	87
4.4	<b>Considerações Parciais</b> . . . . .	90
5	<b>RESULTADOS</b> . . . . .	91
5.1	<b>Rastreabilidade dos Atributos Seleccionados para Composição da Árvore de Decisão</b> . . . . .	91
5.2	<b>Regras Extraídas pela DT Proposta para Análise de Gases Dissolvidos no óleo de Transformadores de Força</b> . . . . .	104
5.3	<b>Uma Nova Proposta para o Triângulo de Duval Utilizando Árvores de Decisão</b> . . . . .	107
5.4	<b>A Regra do Acetileno para classificação DGA</b> . . . . .	111
5.5	<b>Utilização da DT Proposta e da Regra do Acetileno em DGA</b> . . . . .	113
5.5.1	<i>Casos encontrados em Siada e Hmood (2013)</i> . . . . .	113
5.5.2	<i>Casos encontrados em Wang (2014)</i> . . . . .	114
5.6	<b>Considerações Parciais</b> . . . . .	118
6	<b>CONCLUSÕES</b> . . . . .	119
6.1	<b>Trabalhos Futuros</b> . . . . .	120

## 1 INTRODUÇÃO

A atual busca pela otimização no ciclo de vida dos ativos inseridos no sistema elétrico de potência, motiva diariamente o desenvolvimento de pesquisas voltadas para a descoberta de novas metodologias para diagnóstico e tratamento de falhas em equipamentos previamente diagnosticadas. O transformador de força destaca-se como sendo o equipamento de maior custo e importância dentro de uma subestação de energia elétrica, devendo portanto estar disponível ininterruptamente durante todo seu ciclo de operação.

Ultimamente, pesquisas tem se empenhado no desenvolvimento de métodos para o correto diagnóstico de falhas incipientes em transformadores de força como uma ferramenta a ser utilizada nos processos de manutenção preditiva e preventiva deste equipamento.

Na seção 1.1 serão apresentadas as motivações para o desenvolvimento da pesquisa a ser apresentada neste trabalho, bem como uma contextualização atual da problemática abordada. Na seção 1.2 estão descritos os objetivos a que se propôs a pesquisa e na seção 1.3 será apresentado a forma com que a presente dissertação foi organizada.

### 1.1 Motivação para o desenvolvimento do tema abordado

#### 1.1.1 *O Contexto Atual*

O sistema elétrico de potência brasileiro, segundo dados da Agência Nacional de Energia Elétrica – ANEEL, possui aproximadamente 156 GW de potência instalada subdivididas em Usinas hidrelétricas, termelétricas, termonucleares e eólicas. Essas usinas são responsáveis pela totalidade da geração atual de energia elétrica, contando ainda com uma potência outorgada em construção de 10,7 GW, em que se destaca a energia eólica como a de maior significância no crescimento da matriz.

Toda essa energia gerada deve ser distribuída aos sistemas interligados para que, através de linhas de transmissão e transformadores de potência, possam chegar efetivamente a outros subsistemas de distribuição de energia ou entregues a consumidores finais. Destaca-se nesse contexto a importância do equipamento de transformação apto a operar de forma satisfatória ininterruptamente, afim de evitar penalizações, devido a descontinuidades no fornecimento de energia provocado por faltas, sejam elas permanentes ou desligamentos intempestivos.

Juntando-se ao fato de que a maioria dos transformadores de força em operação

atualmente nas empresas distribuidoras de energia elétrica no Brasil tem o ciclo de vida útil estimado entre 20 e 30 anos, a dificuldade de fabricação imediata de equipamento de tal porte, grande importância vem sendo dada às técnicas de diagnósticos de falhas internas em transformadores, apresentadas no corpo das normas (IEEE, 2008; IEC, 2008) como principal ferramenta de manutenção preventiva e preditiva adotada por empresas concessionárias transmissoras e distribuidoras de energia elétrica. Essas técnicas buscam a detecção precoce de falhas que podem, em um quadro evolutivo, provocar avarias permanentes e irreversíveis em transformadores de força.

Atualmente, o triângulo de Duval (DUVAL E PABLO, 2001; DUVAL, 2002; DUVAL, 2003; DUVAL, 2006) é uma das técnicas de Análise de Gases Dissolvidos (DGA) mais utilizadas e constante no corpo das normas referente a falhas em transformadores de potência. Embora muitos esforços venham sendo feitos para aperfeiçoar essas técnicas, com o uso de métodos provenientes da inteligência computacional, tais como as redes neurais (PATEL and KHUBCHANDANI, 2004; ZHANG et al., 1996; GUARDADO et al., 2001; NARESH, SHARMA e VASHISTH, 2008; BARBOSA, 2008; BARBOSA, 2012; SOUSA et al., 2012). e os sistemas fuzzy (CHEIM et al., 1999; HUANG, YANG and HUANG, 2002 ; WANG, LIU e GRIFFIN, 1998; AMORA et al., 2009; MIRANDA e CASTRO, 2005; HOOSHMAND et al., 2012), a boa aceitação do método proposto por Duval, se dá pelo fato de ser uma técnica de fácil compreensão, visto que as definições das falhas são apresentadas como coordenadas triangulares representadas no próprio triângulo, determinando a falta por regiões pré-definidas do espaço tridimensional.

Embora simples e muito difundida, a metodologia do triângulo de Duval utiliza apenas 03 dos 06 tipos de gases combustíveis que podem ser gerados na ocasião de falta em transformadores, o que representa uma desvantagem para análise devido a perda de informações contidas nos demais gases. Os gases utilizados pelo método do triângulo de Duval são metano (CH<sub>4</sub>), etileno (C<sub>2</sub>H<sub>4</sub>) e acetileno (C<sub>2</sub>H<sub>2</sub>).

Para superar as deficiências relativas à metodologia do triângulo de Duval no diagnóstico de falhas incipientes em transformadores de potência, quanto ao uso de número limitado de gases e ampliar suas regiões de diagnóstico, algoritmos de árvores de decisão podem ser utilizados.

O conceito de Árvore de Decisão (DT, do inglês *Decision Tree*) foi introduzido na década de 90 em QUINLAN (1986), como sendo um modelo de regras que proporcionam um

aprendizado indutivo e supervisionado para a classificação de dados em conjuntos e a predição de séries temporais.

A técnica de DT vem se configurando como uma poderosa ferramenta de mineração de dados, que para atividades de classificação utiliza técnicas para a máxima extração de informações contidas nos dados disponíveis, buscando a seleção dos atributos mais significativos de determinada amostra para posterior divisão e classificação das mesmas.

### ***1.1.2 Motivação para a Pesquisa***

Utilizar uma ferramenta para o diagnóstico de falhas incipientes em transformadores de força com base na análise de gases dissolvidos no óleo isolante do transformador, seja este um método clássico, ou uma ferramenta de Inteligência Computacional, não deve ser tomado como uma simples tarefa de classificação.

Academicamente, utilizam-se diversas bases de dados para treinamento e testes utilizadas em metodologias propostas para classificação de amostras. Porém, classificar o correto tipo de falha que pode estar ocorrendo internamente a um transformador é uma tarefa de alta relevância. Esta relevância deve-se ao alto custo do equipamento em questão, a não existência de sobressalentes para pronta entrega e principalmente os prejuízos advindos da parada repentina do transformador, gerando prejuízos tanto para as empresas concessionárias de energia elétrica quanto para os consumidores.

Nesse contexto, é de fundamental importância para empresas concessionárias de energia e seus departamentos de manutenção, um minucioso cuidado com o parque de transformadores de força, utilizando todos os métodos disponíveis para monitoramento, acompanhamento e diagnóstico das possíveis falhas que podem ocorrer internamente a transformadores de força em operação.

Buscando uma alternativa que venha a otimizar os métodos atualmente utilizados para realizar a análise de gases dissolvidos no óleo de transformadores, servindo de alternativa a ser utilizada em conjunto com os já empregados, este trabalho propõe o desenvolvimento de metodologia de análise de falhas incipientes em transformadores com a utilização de árvores de decisão.

## 1.2 Objetivos da Pesquisa

O presente trabalho, faz referência aos métodos utilizados para análise de gases dissolvidos extraídos das amostras do óleo de transformadores em operação. Métodos tradicionais, apresentam lacunas (regiões de indefinição) no momento da classificação, bem como a utilização de parâmetros que dependem da presença de determinado gás para validação do método.

Métodos atuais, baseados em Inteligência Computacional, alcançam um elevado percentual de assertividade, porém, podem utilizar algoritmos complexos, procedimentos de classificação que podem não ser claros para o especialista e em alguns casos uma mescla de todas as possibilidades para emissão de determinado diagnóstico.

O objetivo da presente pesquisa é a extração de regras para melhorar o desempenho de DGA em comparação aos métodos clássicos, melhorando o grau de assertividade alcançado pelos métodos atuais utilizando árvores de decisão.

Os objetivos específicos são:

- Induzir árvores de decisão utilizando o algoritmo C4.5, comparando-as até a escolha da melhor opção para classificação DGA;
- Validar as proposições apresentadas pela árvore de decisão final, acerca dos gases utilizados para classificação DGA com a teoria de geração dos gases (HALSTEAD, 1973);
- A concepção de uma metodologia de alto desempenho utilizando regras extraídas da árvore de decisão proposta para a identificação de falhas incipientes em transformadores de força imersos em óleo, a partir de dados levantados da análise cromatográfica a ser realizada no equipamento;
- Revisitar o método clássico proposto por Duval na concepção do seu triângulo, propondo uma nova perspectiva utilizando DT;
- Propor regra para o diagnóstico de faltas incipientes em transformadores divididas em térmicas e elétricas.

Para alcançar os objetivos propostos pela pesquisa, utilizar-se-á o algoritmo de indução de árvores C4.5 para a formatação de árvores de decisão para classificação de falhas em transformadores. Após a criação das árvores será escolhida a de melhor assertividade na classificação das amostras. Deve-se elaborar um estudo detalhado acerca de cada um dos atributos a serem utilizados para a classificação DGA, relacionando assim, os atributos que estão envolvidos quando da ocorrência de determinado tipo de falha.

Posteriormente, deve-se validar fisicamente os resultados extraídos da árvore de decisão final utilizando a comparação dos gases apresentados nos ramos da árvore com a teoria de geração de gases (HALSTEAD, 1973).

O uso de árvores de decisão como ferramenta de pesquisa, deve-se ao fato de que as informações necessárias para a composição dos nós das árvores, são extraídas dos atributos através de métricas utilizadas em teoria da informação como Entropia e Razão de Ganho, que são parâmetros que efetivamente medem a contribuição de um determinado grupo de atributos para a composição da máxima informação.

Conforme citado acima, também é objetivo da Dissertação a utilização de algoritmos de indução de Árvores de Decisão para aumentar a performance do método clássico do triângulo de Duval, uma vez que este, devido a sua simplicidade e satisfatório percentual de classificação, configura-se como um método de boa aceitação atual.

A contribuição ao final do desenvolvimento da referida pesquisa está na concepção de metodologias para utilização em empresas concessionárias de energia para o auxílio da gestão da manutenção de transformadores de força, no que tange a análise de falhas incipientes, bem como um material de apoio que norteie o diagnóstico rápido e preciso da ocorrência de faltas em transformadores de força em operação.

### **1.3 Disposição Geral do documento**

A presente dissertação está disposta em seis capítulos. No capítulo 1 faz-se a contextualização da temática abordada no presente trabalho à situação atual do sistema elétrico de potência, apresentando a motivação para o desenvolvimento da pesquisa.

No capítulo 2, é feita abordagem qualitativa dos tipos de óleos utilizados em transformadores de força, formação e processos de degradação. Também serão abordados neste capítulo, os tipos de falhas incipientes em transformadores e os métodos de análises dos gases dissolvidos, dando conotação adicional aos métodos clássicos utilizados pelas normas. Por fim, serão apresentadas as limitações dos métodos clássicos de análise, bem como as propostas atuais para o tema.

No capítulo 3, são apresentados os conceitos acerca de *Data Mining* e do processo de descoberta e extração de conhecimento em base de dados. Serão desenvolvidas neste capítulo, as bases para a utilização da metodologia proposta que são as árvores de decisão. Finalizando o mesmo, serão apresentados dois algoritmos de indução de árvores, que são o ID3 e o C4.5, este

último, utilizado no presente trabalho para o desenvolvimento da metodologia proposta.

No capítulo 4, após revisão bibliográfica acerca da indução de árvores de decisão, essas técnicas são utilizadas para a formatação da árvore de decisão para fins de classificação DGA. A priori, será realizado o aumento do espaço de entrada inicial, fornecendo assim uma opção maior ao algoritmo de escolha dos melhores atributos para a indução da árvore.

No capítulo 5, após já definido o formato da árvore de decisão de melhor resultado, são apresentados os resultados encontrados e propostas novas metodologias alternativas para análise de gases dissolvidos em transformadores.

Por fim, no capítulo 6 são apresentadas as conclusões sobre a pesquisa realizada, bem como as propostas para trabalhos futuros.

#### **1.4 Produção Científica**

A pesquisa realizada neste trabalho resultou em trabalhos científicos, dos quais se destacam:

MENEZES, A. G. da C.; ALMEIDA, O. M.; BARBOSA, F. R. Use of data mining algorithm to extract information contained in chromatographic data for the analysis of dissolved gases in power transformers. *Conference on Compatibility, Power Electronics and Power Engineering (IEEE CPE-POWERENG)*, Doha, 2018.

MENEZES, A. G. da C.; ALMEIDA, O. M.; BARBOSA, F. R. Use of decision trees for extraction of information contained in chromatographic data for analysis of dissolved gases in transformers. *World Congress on Computational Intelligence (WCCI)*, Rio de Janeiro, 2018.

MENEZES, A. G. da C.; ALMEIDA, O. M.; BARBOSA, F. R. Use of decision trees for extraction of information contained in chromatographic data for analysis of dissolved gases in transformers. *International Conference on Intelligent and Advanced System (ICIAS)*, Kuala Lumpur, 2018

MENEZES, A. G. da C.; ALMEIDA, O. M.; BARBOSA, F. R. Use of decision tree algorithms to diagnose incipient faults in power transformers. *Simpósio Brasileiro de Sistemas de Energia (SBSE)*, Niterói, 2018

## **2 FALHAS INCIPIENTES EM TRANSFORMADORES DE FORÇA**

A detecção de gases gerados em um transformador de potência em serviço imerso em óleo mineral é frequentemente a primeira indicação disponível do estado de funcionamento do equipamento que, eventualmente pode levar à falha se não for corrigida a anomalia geradora do gás (IEEE, 2008).

A análise dos gases dissolvidos no óleo isolante é uma técnica bastante utilizada e permite determinar a condição de operação do isolamento, identificando as eventuais falhas que podem ocorrer ou que estão em processo inicial de formação (falhas incipientes) (Lima, 2008).

Algumas condições que ocorrem isoladas ou simultâneas, podem resultar na decomposição dos materiais isolantes e na formação de vários gases combustíveis e não combustíveis (IEEE, 2008).

Eventos como arco elétrico, descargas parciais, descargas elétricas de alta e baixa energia e sobreaquecimentos do óleo ou dos enrolamentos, são alguns dos principais fatores que internamente ao transformador contribuem para a formação de gases.

Na seção 2.1, são apresentadas características acerca do óleo isolante utilizado em transformadores de força. Na seção 2.2, são apontadas a classificação das falhas internas em transformadores em térmicas e elétricas. Na seção 2.3, serão abordados os métodos convencionais de análise DGA em transformadores de força. Na seção 2.4, serão discutidos as limitações existentes nos métodos de análises convencionais e na seção 2.5 serão apresentadas as conclusões finais do capítulo.

### **2.1 O óleo isolante em transformadores de força**

O óleo mineral isolante está em contato com todas as partes ativas internas do transformador, e além da função de isolamento, também acumula a função de meio refrigerante. Devido a essas razões, qualquer alteração no funcionamento do equipamento, seja envelhecimento ou por esforços de trabalho acima das condições de normalidade, reflete-se nas características e no comportamento do meio isolante (BARBOSA, 2013).

A análise do óleo isolante assume papel importante em planos de manutenção de transformadores. Um estudo sobre transformadores de potência e as possíveis falhas que o tirem de operação deve abordar a análise do óleo isolante (BARBOSA, 2013).

As condições de funcionamento e integridade de um transformador imerso em óleo

podem ser estabelecidas a partir da análise do seu óleo isolante. Sabe-se que um conjunto de análises definidas por normas técnicas permite o diagnóstico das condições de falhas do transformador (BARBOSA, 2008).

Os métodos mais estudados de diagnóstico utilizados para identificar falhas incipientes em transformadores de potência são: i) avaliação físico-química, que determina o estado do óleo; ii) análise cromatográfica que verifica possíveis falhas no equipamento. A partir destes dois tipos de análises foram definidas normas e procedimentos de operação e manutenção de equipamentos isolados a óleo. (BARBOSA, 2008)

### **2.1.1 Óleo mineral isolante(OMI)**

O OMI utilizado em equipamentos elétricos é obtido pela destilação do petróleo e corresponde à fração obtida entre 300 e 400 °C , ilustrada na Figura 1. Este destilado pode ser de origem parafínica ou naftênica (LIPSHTEIN e SHAKHNOVICH, 1970 e CLARK, 1962). Trata-se de uma mistura de compostos na qual a maioria das moléculas é constituída por carbono e hidrogênio (hidrocarbonetos) e, em pequenas quantidades, por compostos que apresentam nitrogênio, enxofre e oxigênio em sua estrutura (WILSON, 1980).

Por mais de cem anos, o OMI proveniente do refino do petróleo tem sido usado como líquido isolante em transformadores. Muitos bilhões de litros de óleo estão em uso em equipamentos elétricos pelo mundo. A popularidade do OMI se deve a sua disponibilidade e seu baixo custo. Adicionalmente ele é um excelente meio de isolamento e de resfriamento (CLAIBORNE & CHERRY, 2006).

A classificação do OMI em parafínico ou naftênico é, normalmente, realizada segundo a norma ASTM D 2140, por meio de medidas de viscosidade, densidade, índice de refração e pela correlação entre estes resultados em um diagrama ternário (NYNÄS, 2004).

Alternativamente, a classificação do OMI em parafínico ou naftênico pode ser realizada com a utilização da técnica de espectroscopia na região do infravermelho, pela determinação do seu conteúdo parafínico (CP). Óleos com CP inferiores à 50% são considerados naftênicos enquanto que aqueles com CP iguais ou superiores à 56% são classificados como parafínicos. Os óleos com valores de CP entre 50 e 56% são classificados, em geral, como óleos intermediários (NYNÄS, 2004).

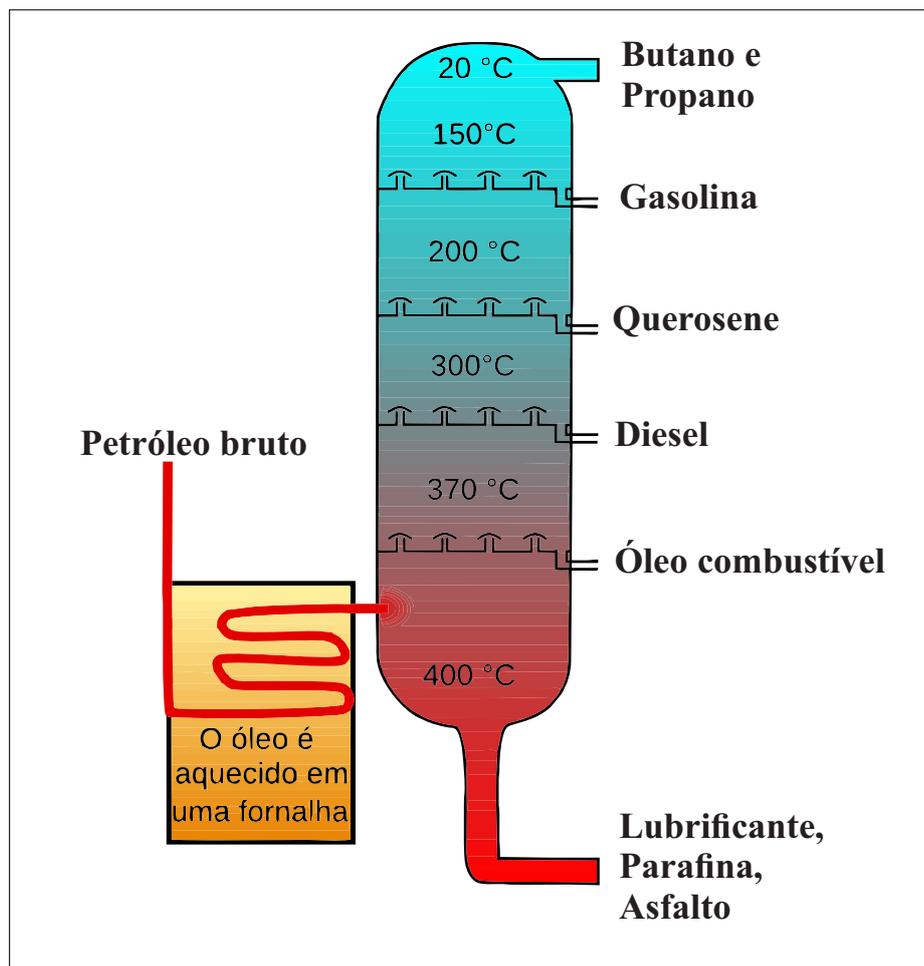
### 2.1.1.1 Óleo mineral isolante naftênico

Este óleo também é conhecido como óleo mineral isolante do tipo A. O óleo mineral isolante naftênico são hidrocarbonetos saturados de cadeia fechada, podendo ter um ou mais anéis com uma ou mais cadeias laterais. Este é o tipo de óleo mais utilizado em transformadores, pois apresenta boa condutividade térmica, alto ponto de fulgor, baixo custo de produção em larga escala, alta capacidade de regeneração, baixa toxicidade e biodegradabilidade. Na figura 2 está representada uma cadeia naftênica (FRIEDENBERG; SANTANA, 2014).

### 2.1.1.2 Óleo mineral isolante parafínico

Os parafínicos, também conhecidos como óleo mineral isolante tipo B, por serem derivados de petróleo, são hidrocarbonetos saturados com cadeia aberta linear ou ramificada. Através das figuras 3 e 4, podemos observar as cadeias parafínicas linear e ramificada (FRIE-

Figura 1 – Destilação fracionada do petróleo



DENBERG; SANTANA, 2014).

### 2.1.2 Comparação entre os óleos parafínicos e naftênicos

Apesar dos dois tipos de óleo virem da mesma origem, algumas características do petróleo proporcionará maiores produções de um ou de outro. Ainda que os modernos processos de refinação possam modificar as características do óleo, esses dois tipos de óleo contém propriedades distintas, fazendo com que cada um seja mais adequado para um determinado fim, entretanto, essas diferenças não fazem de um óleo melhor que o outro (MOURA E CARRETEIRO, 1975).

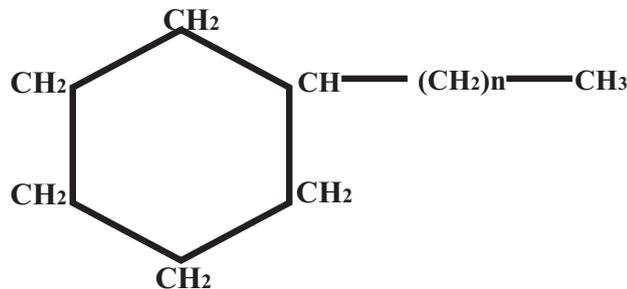
Os óleos parafínicos apresentam uma taxa de oxidação inferior ao naftênicos, entretanto, os produtos da oxidação (borra) são mais solúveis nos óleos naftênicos, evitando que estes produtos da oxidação não se depositem no fundo do tanque do transformador, mantendo o bom funcionamento do sistema de arrefecimento, já que não terá partículas atrapalhando a circulação do óleo por convecção. Outras características distintas podem ser encontradas na Tabela 1 (FRIEDENBERG; SANTANA, 2014).

### 2.1.3 Degradação dos óleos minerais isolantes

Durante a operação do equipamento, o OMI envelhece e mudanças consideráveis ocorrem nas suas propriedades físicas, químicas e elétricas (STOCCO, 2009). Como consequências, ocorre:

- deterioração das propriedades isolantes do óleo;

Figura 2 – Cadeia Naftênica



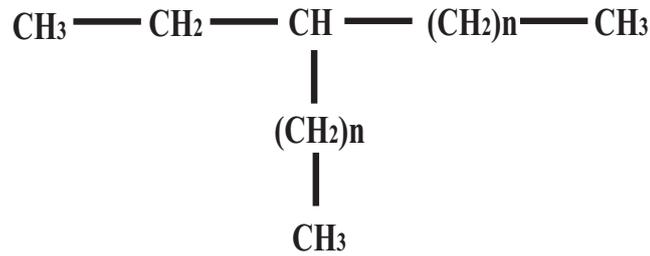
Fonte – FRIEDENBERG; SANTANA, 2014.

Figura 3 – Cadeia Parafínica Linear



Fonte – FRIEDENBERG; SANTANA, 2014.

Figura 4 – Cadeia Parafínica Ramificada



Fonte – FRIEDENBERG; SANTANA, 2014.

Tabela 1 – Características normais dos óleos parafínicos e naftênicos

<b>Característica</b>	<b>Parafínicos</b>	<b>Naftênicos</b>
Ponto de fluidez	Alto	Baixo
Índice de viscosidade	Alto	Baixo
Resistência a oxidação	Grande	Pequena
Oleosidade	Pequena	Grande
Resíduo de carbono	Grande	Pequeno
Emulsibilidade	Pequeno	Grande

Fonte – MOURA E CARRETEIRO

- aceleração do processo de degradação da celulose;
- formação de borra (material que se deposita na parte ativa dos transformadores, dificultando a transferência de calor).

### 2.1.3.1 Processo de oxidação dos óleos minerais isolantes

Os OMIs são constituídos basicamente por compostos químicos da classe dos hidrocarbonetos. Portanto, o seu comportamento com relação à oxidação será determinado por estes compostos que sob a ação do calor reagem com o oxigênio dissolvido no óleo formando os hidroperóxidos.

A presença de um radical livre, o qual pode ser formado fotoquimicamente por ativação térmica, é suficiente para formar inúmeros radicais livres por meio da reação em cadeia, desta forma, a velocidade da reação é uma função exponencial com o tempo a uma dada temperatura, sendo que a velocidade da reação também aumenta com o aumento da temperatura de forma exponencial (STOCCO, 2009).

Após a formação dos hidroperóxidos, muitos produtos de oxidação são formados. Numa etapa seguinte, os produtos (álcoóis, aldeídos, cetonas, chamados de produtos intermediá-

rios da oxidação), na presença de oxigênio, dão origem aos ácidos carboxílicos.

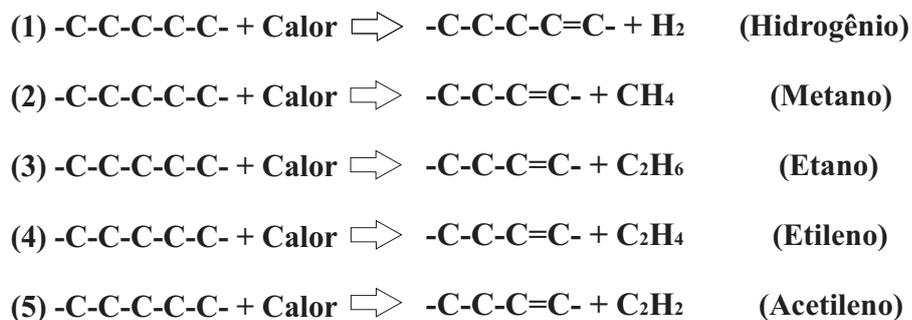
Posteriormente, ocorrem reações de polimerização, ou seja, várias moléculas reagem formando um composto de alta massa molar (mais pesado que o óleo) que se deposita sob a forma de borra. Esta borra pode ser de caráter apolar (polimerização via radical livre) e de caráter polar e ácido (polimerização de ácidos e alcoóis). Esses produtos finais de oxidação são compostos oxigenados com polaridade na molécula e, portanto maus isolantes elétricos. Além disso, os compostos ácidos formados são agentes aceleradores da degradação do papel isolante (MILASCH, 1984).

Os produtos poliméricos formados na última etapa da reação são extremamente prejudiciais à operação do equipamento, já que pela sua alta massa molar são sólidos que se depositam sobre a parte ativa do transformador impedindo a transmissão do calor. Por serem oxigenados, tendem a acumular água, acelerando ainda mais a degradação do papel (MILASCH, 1984).

### 2.1.3.2 Degradação Térmica dos Óleos Minerais

Os hidrocarbonetos pirolizam gerando outros hidrocarbonetos de menor massa molar. Isto é, sob a ação do calor suas estruturas degradam, gerando fragmentos de vários tamanhos, incluindo compostos de massa molar muita baixa e, portanto, gasosos. Esta reação obedece a leis termodinâmicas predeterminadas e assim, os produtos gerados pela pirólise são função da temperatura a que o óleo é submetido. As reações de pirólise mais importantes para a manutenção do OMI são as apresentadas na Figura 5. As reações 1 e 3 ocorrem a baixas temperaturas, em seguida, tem-se as reações 2, 4 e 5 em ordem crescente de temperatura (STOCCO,2009).

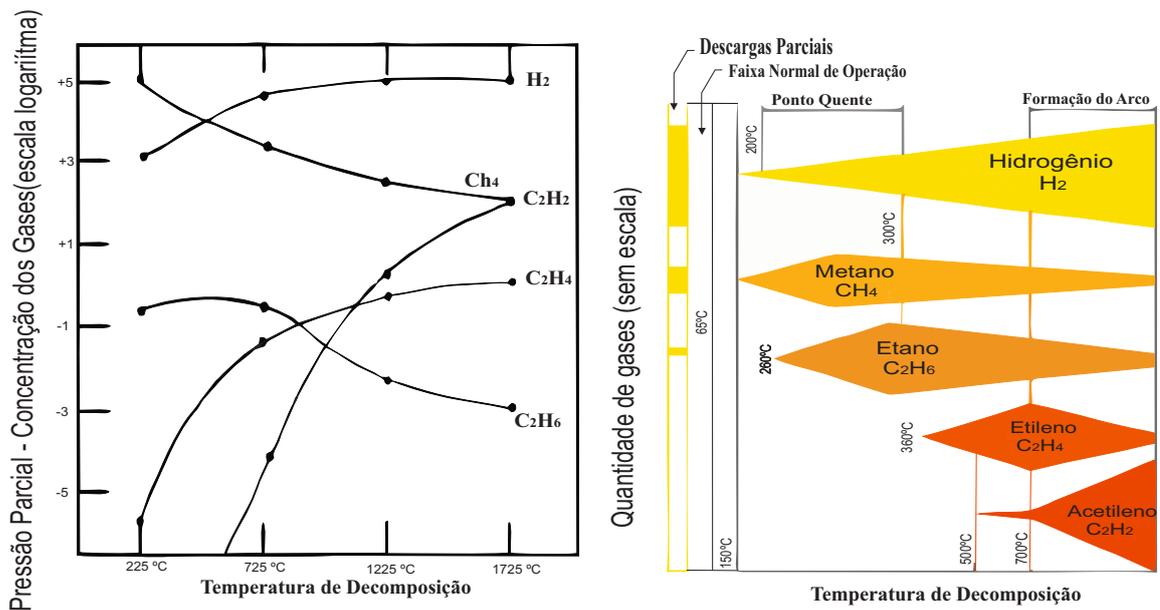
Figura 5 – Reações de Pirólise do OMI.



Fonte – STOCCO, 2009 .

O modelo termodinâmico proposto por Halstead (1973) permite conhecer a pressão parcial de cada gás em função da temperatura (vide Figura 6). Esta proposta admite que todos os hidrocarbonetos do óleo são decompostos nos mesmos produtos e que estes estão em equilíbrio entre si. Idealmente, toma-se que as reações ocorrem em temperatura constante (apud Barbosa, 2013).

Figura 6 – Formação dos gases em função da temperatura de decomposição. a) pressão x temperatura. b) concentração dos gases x temperatura.



Fonte – a) IEEE (2008). b) Adaptado de Barbosa (2013).

Este modelo pode ser utilizado na estimação da taxa de geração de cada gás a qualquer temperatura. Com isso, uma distribuição da concentração de gases pode significar uma determinada falha caracterizada em termos de temperatura e duração da falha (IEEE, 2008).

## 2.2 Falhas em transformadores

O sistema de isolamento de um transformador, mesmo operando em condições de normalidade, sofre redução de sua capacidade à medida que envelhece. No caso de operação em condições críticas, esta redução pode ser acelerada e a vida útil do isolamento se torna mais breve. Quando a capacidade do sistema de isolamento é reduzida, vários processos podem ser desencadeados, acelerando o processo de incapacidade do isolamento (Barbosa, 2013)

Inerente ao princípio de funcionamento de um transformador, ocorrem perdas pro-

duzidas no ferro e no cobre. Estas perdas geram calor que leva ao aumento da temperatura das partes ativas do equipamento. Este aumento de temperatura causa desequilíbrio térmico entre a parte aquecida e o ambiente. No caso de transformadores de potência, este meio é constituído por óleo isolante e papel celulósico para onde é conduzido parte deste calor (ALMEIDA et. al., 2007b).

Quando solicitações térmicas e/ou elétricas, oriundas da operação do transformador de potência, produzem calor acima da capacidade de refrigeração do sistema de isolamento, há um agravamento da sua integridade como isolante e falhas incipientes passam a ocorrer no interior do equipamento. A vida útil do sistema de isolamento de um transformador de potência sofre impacto direto do aumento de temperatura que desencadeia e acelera a degradação do isolamento (ALMEIDA et. al., 2007b).

As falhas incipientes podem ser classificadas em termos dos esforços térmicos, nos quais sobreaquecimentos são atores principais, e elétricos, relacionados às descargas internas no equipamento. As falhas de características térmicas resultam de correntes elétricas que circulam no equipamento e do seu relacionamento com a temperatura ambiente. Os esforços elétricos abrangem tensão e frequência de operação. Estes esforços podem, ainda, ocorrer tanto de forma isolada quanto de forma simultânea (AZEVEDO, 2009).

Adiciona-se a isso a possibilidade do aparecimento de bolhas que contribuem para diminuição da capacidade dielétrica do sistema de isolamento líquido devido a presença de umidade, contaminações do processo de industrialização e impurezas no material (AZEVEDO, 2009).

### **2.2.1 Falhas Elétricas**

As falhas elétricas estão associadas às descargas internas no equipamento, envolvendo tensão e frequência de operação. As falhas elétricas em transformadores de potência são conhecidas como descargas parciais, corona e arco elétrico. Descargas parciais são caracterizadas por uma série de descargas elétricas rápidas por um meio isolante entre duas partes ativamente condutoras. As descargas parciais de baixa intensidade aparecem devido à baixa eficiência dos materiais isolantes, entretanto, não geram deteriorações perceptíveis nos materiais envolvidos. As descargas parciais de alta intensidade surgem em decorrência de sobre-solicitação do isolamento. Neste caso, aparecerão sinais de deterioração dos materiais (BARBOSA, 2013; HELL, 2002).

A descarga elétrica de baixa energia (corona) é ocasionada por um aumento no gradiente da tensão aplicada entre dois condutores que leva o meio isolante em volta dos condutores a se tornar condutor. No ambiente do transformador, o isolante líquido é o meio ionizado. O arco elétrico, ou descarga elétrica de alta energia, apresenta duração efetiva e é caracterizado pelas maiores temperaturas entre as falhas elétricas, ocorrendo devido a curtos-circuitos internos dos enrolamentos, ou curtos-circuitos em conexões de potenciais contra a terra ou devido a descargas entre condutores não-isolados (BARBOSA, 2013; HELL, 2002).

Para as descargas elétricas de baixa intensidade de energia, a produção de hidrogênio apresenta-se elevada, com pequena quantidade de metano e apenas possíveis traços de acetileno. À medida que a intensidade de energia das descargas aumenta, e, conseqüentemente, aumenta a temperatura na vizinhança da falha, ocorre uma significativa aceleração da produção de etileno e acetileno. Se a quantidade de gases dissolvidos é elevada e ocorrer uma alteração na temperatura que modifique o valor da solubilidade de saturação, os gases antes dissolvidos podem ser liberados na forma de bolhas. A existência de bolhas de gases não dissolvidos altera a condição de rigidez dielétrica, levando a uma ruptura do isolamento (LIMA, 2005).

No caso da intensidade das descargas elétricas atingirem valores de alta energia, ou seja, a formação de descargas contínuas, como arcos elétricos, a produção de acetileno torna-se relevante, uma vez que a temperatura se eleva entre 800°C e 2800°C. Possíveis causas de formação destes pontos quentes são: centelhamentos promovidos por más conexões ou interrupção de correntes no comutador, e descargas de contorno ou arcos de potência decorrentes da ruptura dielétrica do óleo entre os enrolamentos, entre espiras ou entre espiras e massa (LIMA, 2005; BARBOSA, 2008).

As Figuras 7 e 8 ilustram respectivamente a parte ativa (vide Figura 7) de um transformador de força 69/13.8 kV com a potência de 25 MVA que apresentou avaria após a ocorrência de uma falha elétrica, bem como o local de ocorrência na falha (vide Figura 8) na parte inferior do enrolamento da fase B do referido equipamento.

### **2.2.2 Falhas Térmicas**

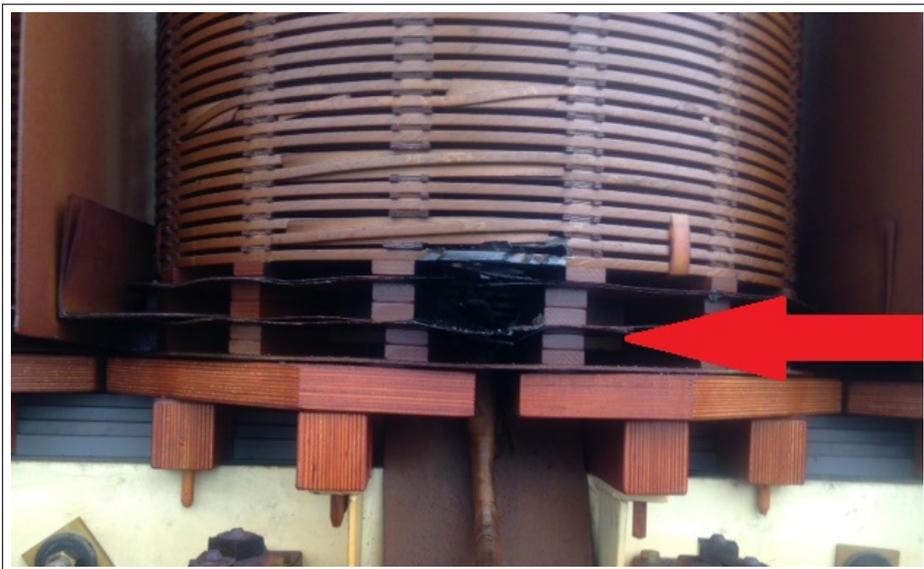
As falhas térmicas são caracterizadas por sobreaquecimentos. Os sobreaquecimentos de até 300°C, que podem ser detectados nos enrolamentos ou pontos quentes no núcleo, são ocasionados por refrigeração insuficiente, sobrecargas ou perdas excessivas de magnetização. Sobreaquecimentos de até 1000 °C são causados por correntes de Foucault no núcleo ou devido

Figura 7 – Transformador de força 69/13.8 kV com a potência de 25 MVA avariado em ocorrência de falha elétrica



Fonte – Eletrobrás Distribuição Piauí, 2015

Figura 8 – Avaria provocada por falha elétrica na parte inferior do enrolamento central do transformador



Fonte – Eletrobrás Distribuição Piauí, 2015

ao mau contato de seletores de tensão. Sobreaquecimentos acima de 1000°C são ocasionados por curtos-circuitos interlaminares no núcleo (HELL, 2002).

O sobreaquecimento do óleo isolante induz a produção de gases etileno ( $C_2H_4$ ) e metano ( $CH_4$ ), juntamente com pequenas quantidades dos gases hidrogênio ( $H_2$ ) e etano ( $C_2H_6$ ). Se o sobreaquecimento resultar em temperaturas muito elevadas podem ser detectadas pequenas quantidades de acetileno ( $C_2H_2$ ) (ARANTES, 2005; GOMEZ, 2013).

A decomposição térmica da celulose e outros isolamentos sólidos produzem óxidos de carbono (CO e  $CO_2$ ) em quantidades relevantes já em temperaturas muito baixas. Portanto, a relação entre monóxido (CO) e dióxido de carbono ( $CO_2$ ) carrega em si a importância de indicar a decomposição térmica da celulose em termos quantitativos.

Sendo assim, o IEEE Std. C57.104-2008 (IEEE, 2008) utiliza-se da razão  $CO_2/CO$  e estabelece que sete (7,0) é um valor típico, muito embora alguns pesquisadores contestem (KAN e MIYAMOTO, 1995; IEEE, 2008). De forma mais geral, uma razão  $CO_2/CO$  na faixa de 3,0 a 10,0 é considerada normal. Também, de modo semelhante, a norma IEC 60599 (IEC, 1999) propõe a avaliação da degradação da celulose através do monitoramento dos óxidos de carbono, no entanto, utiliza a razão  $CO/CO_2$ , sendo definida uma faixa de valores ( $0,07 < CO/CO_2 < 0,30$ ) para uma condição normal de envelhecimento (WANG, 2000; BARBOSA, 2013).

Em uma situação da celulose sob temperatura extremamente alta, como na ocorrência de arcos elétricos, a razão  $CO/CO_2$  aproxima-se de 1, pois a geração de monóxido de carbono (CO) aumenta muito rapidamente em comparação à produção de dióxido de carbono ( $CO_2$ ). Porém, em uma situação de leve sobrecarga ou problemas de ventilação, onde ocorre leve sobreaquecimento, o  $CO_2$  cresce muito mais rapidamente que o CO, portanto a taxa de  $CO/CO_2$  fica na faixa entre 1:10 e 1:20 (LIMA, 2005; BARBOSA, 2008).

Essas ocorrências de falhas, sejam térmicas ou elétricas, resultam na formação de gases que são parcial ou totalmente dissolvidos no óleo, sendo assim transportados para todos os pontos que o óleo possa atingir. Desse modo, através de análises de amostras do óleo é possível obter informações de todos os componentes internos em contato com o mesmo (NOGUEIRA, 2004).

## 2.3 Métodos convencionais utilizados para análise de gases dissolvidos no óleo do transformador de força

Transformadores de potência imersos em óleo mineral sob condições normais de operação geram gases combustíveis em baixas concentrações e diretamente proporcionais ao seu tempo de utilização, degradação natural da celulose e do próprio óleo.

Durante a ocorrência de uma falta, operação anormal ou fora dos parâmetros nominais do transformador, aumentam a quantidade de gases combustíveis produzidos através da recombinação química dos elementos desprendidos das cadeias de hidrocarbonetos que compõem o óleo.

Em baixas temperaturas inicia-se a formação de  $H_2$ , que caracteriza a existência de descargas elétricas de baixa energia no óleo do transformador. Posteriormente, em decorrência do progressivo aumento da temperatura surgirão  $CH_4$ ,  $C_2H_6$ ,  $C_2H_4$  e  $C_2H_2$ , sendo a presença de acetileno um forte indicativo para a ocorrência de arco elétrico no óleo do transformador, podendo atingir temperaturas próximas a  $1000^\circ\text{C}$ .

Dentre os procedimentos descritos na norma IEEE C57.104 e IEC 60599, cita-se o método do Gás chave, o método das razões de Dornenburg e o triângulo de Duval. Todos estes métodos têm sido aplicados atualmente e apresentam boa aceitação.

### 2.3.1 Análise do Gás Chave

O método do gás chave faz uso do valor percentual dos gases para diagnosticar falhas incipientes em transformadores. A essência do método considera o conhecimento da temperatura na qual um gás é predominantemente gerado e a correlação dessa temperatura, e consequentemente do gás, a uma determinada falha. O gás predominante é o gás chave ou principal (LIMA, 2005).

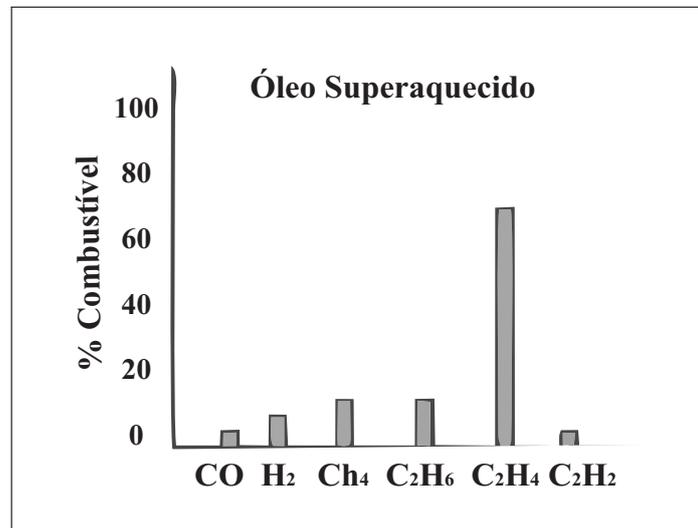
Abaixo, tem-se a distribuição percentual de cada concentração de gás combustível, associada ao tipo de falha pelo método de citado.

#### 2.3.1.1 Falha térmica no óleo

Produtos da decomposição incluem etileno ou eteno ( $C_2H_4$  – cerca de 63%) e metano ( $CH_4$  – 16%), junto com menor quantidade de hidrogênio  $H_2$  – 2%) e etano ( $C_2H_6$  – 19%). Traços de acetileno ( $C_2H_2$ ) podem ser formados se a falha for severa ou envolver contatos elétricos. O

gás principal indicativo, como mostrado pelas proporções, é o etileno (IEEE,2008).

Figura 9 – Diagnóstico: Falha térmica no óleo



Fonte – IEEE, 2008

#### 2.3.1.2 Falha térmica na celulose

Grandes quantidades de monóxido de carbono (CO – 92%) e dióxido de carbono (CO<sub>2</sub>) são desenvolvidas do sobreaquecimento da celulose. Gases hidrocarbonetos, tais como etileno e metano, serão formados se a falha envolve uma estrutura impregnada de óleo. O gás indicativo é o monóxido de carbono (IEEE,2008).

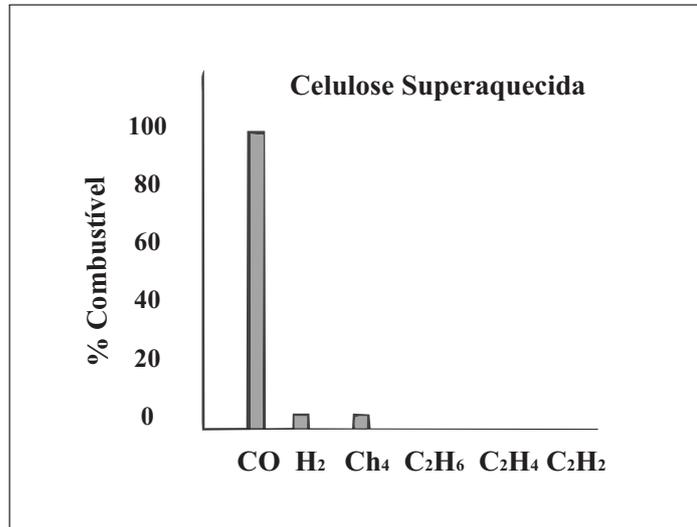
#### 2.3.1.3 Descarga parcial ou Corona

Produz hidrogênio (H<sub>2</sub> - 85%) e metano (CH<sub>4</sub> - 13%), com pequenas quantidades de etano e etileno. Quantidades comparáveis de monóxido e dióxido de carbono podem resultar de descargas na celulose. O gás indicativo é o hidrogênio (IEEE, 2008).

#### 2.3.1.4 Descarga de Alta Energia ou Arco

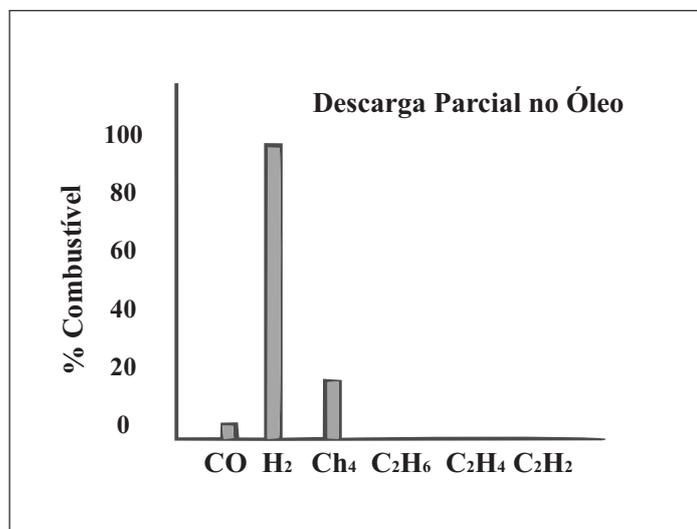
Grandes quantidades de hidrogênio (H<sub>2</sub> - 60%) e acetileno (C<sub>2</sub>H<sub>2</sub> - 30%) são produzidas, com menores quantidades de metano (CH<sub>4</sub> - 5%) e etileno (C<sub>2</sub>H<sub>4</sub> - 3%). Dióxido e

Figura 10 – Diagnóstico: Falha térmica na celulose



Fonte – IEEE,2008

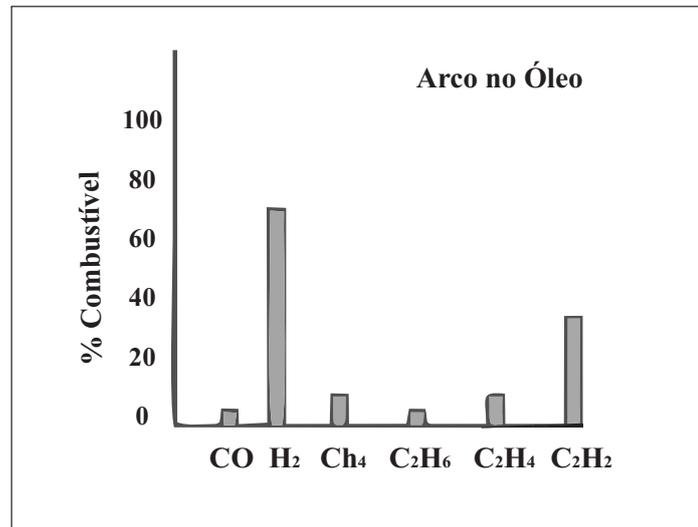
Figura 11 – Diagnóstico: Descarga parcial ou Corona



Fonte – IEEE,2008

monóxido de carbono podem também ser formados se a falha envolve a celulose. Óleo pode ser carbonizado. O gás indicativo é o acetileno (IEEE,2008).

Figura 12 – Diagnóstico: Descarga de alta energia



Fonte – IEEE, 2008

### 2.3.2 Razões de Dornenburg

O uso de razões entre as concentrações dos gases (vide Tabela 2) para indicar um tipo possível de falha é um processo empírico baseado na experiência de cada pesquisador de forma individual, correlacionando as análises de gases de muitas unidades com o tipo de falha posteriormente determinada como a causa para o distúrbio ou falha quando a unidade foi examinada. Esse processo foi atribuído a Dörnenburg e posteriormente confirmado por Rogers em sistemas europeus, dos quais a maior parte da correlação do diagnóstico é obtida (IEEE, 2008)

O método de Dornenburg (1967) não define critérios para o estabelecimento de condições normais do isolamento, mas estabelece as condições de validade para aplicação do método de diagnóstico em função das concentrações individuais dos gases componentes das relações (IEEE, 1991).

O método de Dornenburg, apresentado na Tabela 3 utiliza em sua análise cinco tipos de gases (H<sub>2</sub>, CH<sub>4</sub>, C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, C<sub>2</sub>H<sub>2</sub>) e quatro das relações apresentadas na Tabela 2 para diagnosticar o tipo de falha internas em transformadores como sendo sobreaquecimento,

Tabela 2 – Relação entre gases para DGA

<b>R1</b>	<b>R2</b>	<b>R3</b>	<b>R4</b>	<b>R5</b>
CH <sub>4</sub> /H <sub>2</sub>	C <sub>2</sub> H <sub>2</sub> /C <sub>2</sub> H <sub>4</sub>	C <sub>2</sub> H <sub>2</sub> /CH <sub>4</sub>	C <sub>2</sub> H <sub>6</sub> /C <sub>2</sub> H <sub>2</sub>	C <sub>2</sub> H <sub>4</sub> /C <sub>2</sub> H <sub>6</sub>

Fonte – Randy, 1977

descargas de baixa energia (corona) e descargas de alta energia (arco elétrico).

Tabela 3 – Diagnostico pelo método de Dornenburg

<b>Falha Incipiente</b>	<b>R1</b>	<b>R2</b>	<b>R3</b>	<b>R4</b>
Decomposição Térmica	> 1.0	< 0.75	< 0.3	> 0.4
Corona (Descarga de baixa energia)	< 0.1	Não significativa	< 0.3	> 0.4
Arco (Descarga de Alta Energia)	> 0.1 e < 1.0	> 0.75	> 0.3	< 0.4

Fonte – Randy, 1977

Entretanto, o critério para a validade da utilização dos métodos que utilizam razões é de que um dos gases contidos nas relações principais (CH<sub>4</sub>/H<sub>2</sub> e C<sub>2</sub>H<sub>2</sub>/C<sub>2</sub>H<sub>4</sub>) tenham uma concentração superior ao dobro das quantidades indicada na Tabela 4, e que para as demais relações um dos gases tenha a concentração superior a apresentada na mesma tabela.

Tabela 4 – Limites mínimos para validação do método de Dornenburg

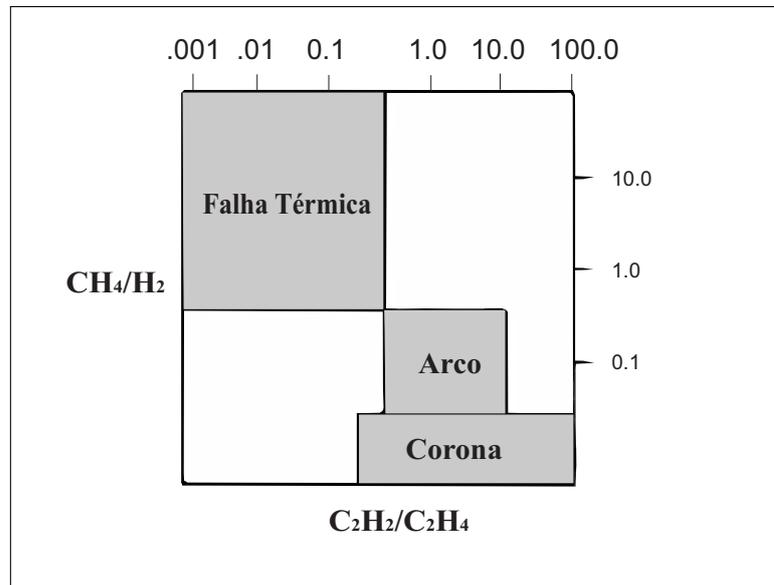
<b>Tipo de Gás</b>	<b>H<sub>2</sub></b>	<b>CH<sub>4</sub></b>	<b>C<sub>2</sub>H<sub>6</sub></b>	<b>C<sub>2</sub>H<sub>4</sub></b>	<b>C<sub>2</sub>H<sub>2</sub></b>
Concentração em ppm	100	120	65	50	35

Fonte – Randy, 1977

O método de Dörnenburg também pode fornecer o diagnóstico por meio da interpretação de um gráfico, em que duas razões de gases são colocadas em eixos e indicam o tipo de falha, conforme indicado na Figura 13. A região de localização dos pontos inseridos é indicativa do tipo de falha do equipamento (BARBOSA, 2013; GOMEZ, 2013).

Como pode ser observado, existem regiões em branco nas quais não se associa nenhum diagnóstico. Logo, este método leva a situações de não-decisão, ou seja, o diagnóstico não poderá ser realizado de modo satisfatório (BARBOSA, 2013).

Figura 13 – Gráfico de Dornenburg



Fonte – Barbosa, 2013

### 2.3.3 Razões de Rogers

O método de Rogers baseia-se nas relações das concentrações dos gases  $H_2$ ,  $CH_4$ ,  $C_2H_6$  e  $C_2H_4$  e nos limites de variações destas relações para emitir o diagnóstico. Este conjunto de relações é constantemente atualizado com o objetivo de ajustar o método aos resultados de novas análises de falhas de equipamentos. É uma das técnicas mais utilizadas nas empresas concessionárias de energia (GOMEZ, 2013).

Entretanto, em caso de ocorrências simultâneas de falhas o diagnóstico pode ser ambíguo, uma vez que as faixas de variação das relações apresentam superposições permitindo mais de uma interpretação (GOMEZ, 2013; SENNA, 2010).

Os primeiros trabalhos de Rogers utilizavam quatro relações de gases. Posteriormente, as normas adotaram o método de Rogers com apenas três relações (vide Tabela 5) de concentração de gases e cinco gases, sendo eliminada a relação  $C_2H_6/C_2H_2$  que era utilizada para identificação de sobreaquecimento de baixa temperatura (ZIRBES, ROLIM, e ZÜRN, 2005).

É possível observar na tabela 05 que as faixas de variação das relações apresentam superposições, não permitindo uma interpretação única. Também em muitos casos não existe um diagnóstico definido, pois as combinações de variações que podem ocorrer nas relações são maiores que aquelas definidas na tabela do método. Esse fato ocorre com todas as variantes do método (ZIRBES, ROLIM, e ZÜRN, 2005).

Tabela 5 – Diagnostico pelo método de Rogers

<b>Diagnóstico</b>	<b>R1</b>	<b>R2</b>	<b>R5</b>
Unidade Normal	$> 0.1e < 1.0$	$< 0.1$	$< 0.1$
Descarga de baixa energia	$< 0.1$	$< 0.1$	$< 0.1$
Descarga de alta energia	$> 0.1e < 1.0$	$> 0.1e < 3.0$	$> 0.3$
Falha térmica com baixas temperaturas	$> 0.1e < 1.0$	$< 0.1$	$> 0.1e < 3.0$
Falhas térmicas $T < 700^{\circ}\text{C}$	$> 1.0$	$< 0.1$	$> 0.1e < 3.0$
Falhas térmicas $T > 700^{\circ}\text{C}$	$> 1.0$	$< 0.1$	$> 0.3$

Fonte – IEEE, 2008

### 2.3.4 Triângulo de Duval

O método de Duval realiza o diagnóstico a partir de um gráfico ternário dividido em áreas de tipos de falhas. Neste método são consideradas as concentrações relativas de três gases combustíveis, metano, acetileno e etileno, conforme indicado nas Equações 2.1, 2.2 e 2.3, sendo que as concentrações dos gases devem estar em partes por milhão (ppm) (DUVAL, 2002).

$$\%CH_4 = \frac{100.[CH_4]}{[CH_4] + [C_2H_2] + [C_2H_4]} \quad (2.1)$$

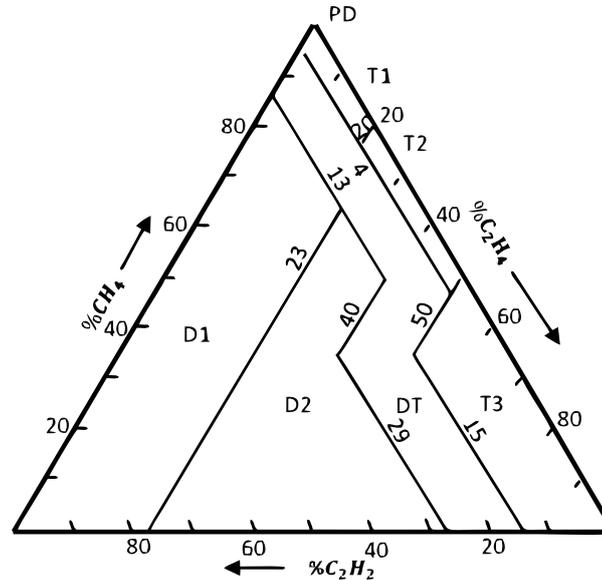
$$\%C_2H_2 = \frac{100.[C_2H_2]}{[CH_4] + [C_2H_2] + [C_2H_4]} \quad (2.2)$$

$$\%C_2H_4 = \frac{100.[C_2H_4]}{[CH_4] + [C_2H_2] + [C_2H_4]} \quad (2.3)$$

Uma vez detectado a existência da falta pelo mapeamento dos gases relacionados a partir das equações acima, utiliza-se o diagrama ternário para a alocação do ponto nas regiões pré-definidas por DUVAL, localizando a que tipo de falta é pertinente a situação analisada (vide Figura 16).

Duval classifica as faltas mapeadas no seu método em descargas parciais (PD), faltas elétricas de baixa (D1) e alta energia (D2), falhas térmicas abaixo de  $300^{\circ}\text{C}$  (T1), falhas térmicas

Figura 14 – Triângulo de Duval



Fonte – Duval, 2002

entre 300°C e 700°C (T2) e falhas térmicas acima de 700°C (T3). Observa-se no triângulo a região definida como DT, atribuídas a faltas de características térmicas e elétricas. Normalmente a classificação de amostras de óleo contidas nessa zona é utilizada para transformadores que contenham comutadores em carga do tipo On Load Tap Changer (OLTC).

Os gases  $C_2H_2$  e  $C_2H_4$  são utilizados em todos os métodos de interpretação para representar falhas de alta energia (tais como arcos) e falhas de alta temperatura.  $H_2$  é preferido em vários destes métodos para representar falhas de energia muito baixa, tais como PD, onde ele é produzido em grandes quantidades.  $CH_4$  foi escolhido para o triângulo, porque não só permite identificar essas falhas, mas oferece melhores resultados globais de diagnósticos para todos os outros tipos de falhas do que quando se utiliza o  $H_2$ . (AMORA, 2012)

## 2.4 Limitações impostas pelos métodos de análise de gases dissolvidos convencionais

As técnicas convencionais de diagnóstico de falhas centradas na análise de gás dissolvido são baseadas mais em evidências empíricas do que em teorias científicas. Portanto, deve ser admitido que a análise de gases e a interpretação dos seus resultados estão sujeitas a variabilidade. Diversos fatores influem para imprecisão das medidas e das interpretações. Entre os motivos para gerar incertezas nas técnicas de análise de gás dissolvido estão os diferentes métodos de fabricação dos equipamentos, dos sistemas de preservação e circulação do óleo, além

do histórico do carregamento dos transformadores (MORAIS, 2004; SENNA, 2010).

Os métodos convencionais de diagnóstico são concebidos para diagnósticos individuais, entretanto, podem ocorrer situações onde estes processos atuam de forma conjunta ou como um processo em cadeia. Situações desse tipo levam a diagnósticos ambíguos ou até mesmo inconclusivos, o que coloca em risco a confiabilidade e a integridade do equipamento (BARBOSA, 2013; SENNA, 2010).

Para a emissão de um laudo técnico há a necessidade de um especialista na área, pois é necessária uma grande experiência e conhecimento das técnicas de análise. Além disso, o especialista não se limita a apenas uma metodologia, pois a dinâmica dos processos envolvidos torna cada caso ímpar. Portanto, os diagnósticos baseados na análise de gases dissolvidos ainda são dependentes da experiência dos especialistas humanos, permanecendo muitas vezes insatisfatórios e incompletos (BARBOSA, 2013).

A base das técnicas de análise de gases dissolvidos é o processo de cromatografia laboratorial, a qual a amostra do óleo isolante do transformador é submetida. Esse processo tem como produto resultante uma estratificação das concentrações dos gases dissolvidos no óleo e, a partir desses valores quantificadores, aplica-se uma técnica de interpretação dos dados. Esses dados podem ser expressos na forma de concentrações individuais de cada gás, através de razões entre as concentrações dos gases ou ainda como porcentagens em relação à concentração total de gases (BARBOSA, 2013).

O processo de falha está intrinsecamente relacionado com as temperaturas esperadas no óleo em função de seus mecanismos. São estabelecidas faixas de temperatura para as quais existem maiores probabilidades da ocorrência de determinadas falhas. Estabelecida a correlação entre a temperatura e a formação de gases no óleo isolante, o processo de falha e sua intensidade podem ser estimados. Essa é a forma usual de estabelecer um diagnóstico do transformador (AMORA, 2013).

A interpretação de uma análise individual pode não ser muito efetiva, pois, mais de uma falha pode estar ocorrendo ao mesmo tempo ou ainda um tipo de falha pode progredir para outro, tal como ocorre com alguns problemas elétricos que têm origem em problemas térmicos. Portanto, é de fundamental importância estabelecer uma base de dados para servir como valores de referência, de forma tal que se torna possível determinar a evolução temporal das concentrações dos gases, tomando-se como base os valores históricos. Isso é implementado por intermédio de coletas periódicas de amostras de óleo, análise cromatográfica para determinar

as concentrações dos gases e formação de um banco de dados com o histórico do transformador (BARBOSA, 2013).

As normas que tratam DGA especificam que quando as relações ou concentrações de gases excedem seus respectivos valores típicos, uma tentativa de diagnóstico pode ser feita aplicando um dos métodos padrões de diagnóstico de análise de gases dissolvidos (métodos de Dörnenburg e Rogers). Esses métodos levam a um número significativo de casos em que nenhum diagnóstico pode ser dado, porque os valores das relações ficam fora da faixa de falhas (LIMA, 2005).

Na análise dos métodos convencionais apresentados pelas normas, observam-se limitações quanto ao uso destes com a finalidade de prover o diagnóstico de falhas incipientes em transformadores de força. No método de Dornenburg, existem regiões em que a falta não é classificada, devido a espaços em branco fora das regiões de decisão. A análise por Rogers pode levar a indefinição do tipo de falta devido a mais de uma possibilidade apontada pelos diagnósticos previstos pelo método.

O triângulo de Duval, com maior aceitação entre os demais métodos para a análise DGA, pode ser penalizado pelo fato de levar em consideração apenas três dos seis gases combustíveis disponíveis para análise, limitando assim a técnica de diagrama ternário utilizado para a representação gráfica do método.

Também, deve ser salientado que nenhum método indicado em normas permite o diagnóstico de falhas simultâneas e nem avaliar diretamente a evolução histórica de uma possível falha. Sendo recomendado ao utilizador, montar um banco de dados histórico com os dados de DGA de um determinado equipamento, de forma a identificar a evolução de falhas ao longo do tempo (AMORA,2013).

## **2.5 Propostas atuais para análise de gases dissolvidos no óleo do transformador de força baseados em Inteligência Computacional**

Nesta seção, são apresentadas algumas ferramentas utilizadas como novas metodologias para o diagnóstico de falhas incipientes em transformadores a partir da análise de gases dissolvidos. São abordadas a utilização de técnicas de inteligência computacional, tais como redes neurais, sistemas de inferência fuzzy e árvores de decisão.

### 2.5.1 *Redes Neurais*

As redes neurais são usadas extensivamente em reconhecimento de padrões. Esta técnica de IA é adaptativa, hábil em tratar relações não-lineares e consegue generalizar soluções para um conjunto novo de dados. Portanto, o diagnóstico de falhas incipientes usando redes neurais pode ser considerado como uma associação entre um conjunto de entradas (concentrações de gases ou razões dos mesmos) e saídas (tipo de falha) (MORAIS, 2004).

Trabalhos recentes empregam redes neurais no diagnóstico do sistema isolante de transformadores de potência e têm alcançado resultados eficientes utilizando técnicas de treinamento adaptativas, uma vez que conseguem lidar de maneira adequada com incertezas de medidas, imprecisões e não linearidades em razão de sua capacidade de aprendizagem (BARBOSA, 2013).

O relacionamento entre os produtos gasosos da degradação do sistema de isolamento de um transformador e as falhas que aceleraram a produção destes gases pode ser realizado por meio de uma RNA. Dessa forma, uma RNA pode ser interpretada como uma ferramenta que fornece alternativas às metodologias convencionais de ensaios, produzindo resultados motivadores, principalmente devido às características intrínsecas da técnica, tais como a sua capacidade de generalização e a facilidade de integração com outras ferramentas computacionais (SILVA, SOUZA e SOUZA, 1999).

Pesquisadores tem utilizado RNA para inferir as relações existentes entre as concentrações e relações entre os gases (entrada da RNA) com o tipo de falha incipiente no transformador (saída da rede)(PATEL e KHUBCHANDANI, 2004; ZHANG et al., 1996; GUARDADO et al., 2001; NARESH, SHARMA e VASHISTH, 2008; SOUSA et al., 2012).

Os sistemas de diagnósticos de falhas incipientes em transformadores utilizando RNA, tornam-se mais robustos uma vez que são utilizados em conjunto com sistemas especialista ou de inferência Fuzzy (CHEIM et al., 1999; HUANG, YANG e HUANG, 2002 ; WANG, LIU e GRIFFIN, 1998; AMORA et al., 2009; MIRANDA e CASTRO, 2005; HOOSHMAND et al., 2012).

Em Guardado (2001) os autores realizaram um estudo comparativo entre diferentes topologias de rede. Foram utilizados cinco métodos para realizar o diagnóstico preliminar dos dados de treinamento: Döernenburg, Rogers, Rogers modificado, IEC e CSUS (IEEE, 1991; IEC, 1999; MYERS, KELLY e PARRISH, 1981). A eficiência das redes ficou entre 87 e 100% e depende do método de diagnóstico utilizado (AMORA, 2013).

Também, em Sousa (2012) foi desenvolvida uma RNA para o diagnóstico de transformadores. Os dados para o treinamento e testes da rede foram baseados no banco de dados gerais de diagnósticos de casos testados da IEC TC 10 (DUVAL e PABLO, 2001), contendo 117 casos. Foram utilizados 70% dos dados para treinamento e 30% para testes da rede neural. A RNA implementada foi do tipo MLP – Multilayer Perceptron (perceptron de múltiplas camadas) (HAYKIN, 2001), com apenas uma camada escondida contendo 10 neurônios e um neurônio na camada de saída. As entradas da rede foram às razões de gases R1, R2 e R5. Para os casos considerados, tanto no treinamento quanto nos testes, a rede neural conseguiu atingir 100% de acerto (AMORA, 2013).

Uma importante vantagem baseada no diagnóstico de faltas utilizando RNA, é a capacidade de aprendizado diretamente a partir de amostras de treinamento, e atualizar seus conhecimentos quando necessário. A alta capacidade de mapeamento não linear dos neurônios garante um desempenho quase sempre superior às soluções dos sistemas fuzzy (ARANTES, 2005).

### ***2.5.2 Sistemas de inferência Fuzzy***

A Lógica Fuzzy é baseada na teoria dos conjuntos Fuzzy. Esta é uma generalização da teoria dos conjuntos tradicionais para resolver os paradoxos gerados a partir da classificação “verdadeiro ou falso” da Lógica Clássica. Tradicionalmente, uma proposição lógica tem dois extremos: ou “completamente verdadeiro” ou “completamente falso”. Entretanto, na Lógica Fuzzy, uma premissa varia em grau de verdade de 0 a 1, o que leva a ser parcialmente verdadeira ou parcialmente falsa, permitindo assim incluir a imprecisão ("fuzziness") nos processos decisórios usando termos lingüístico, tais como: alto, morno, ativo, pequeno, perto, etc. o que caracteriza o grau de pertinência variável (ARANTES, 2005).

Os métodos convencionais de interpretação de análises cromatográficas para diagnosticar a natureza da deterioração do sistema de isolamento de um transformador têm obtido grande aplicabilidade na indústria. Entretanto, tais métodos convencionais falham em alguns casos. Isso acontece, geralmente, quando há mais de uma falha ocorrendo ao mesmo tempo no interior do transformador. Os métodos convencionais definem intervalos fixos para determinadas razões entre as concentrações dos gases, assim, quando uma razão ultrapassa um determinado limiar ocorre uma mudança súbita da faixa à qual ela pertence (0, 1 ou 2). Na realidade, tal transição é suave ao invés de abrupta. Portanto, uma representação fuzzy é capaz de modelar

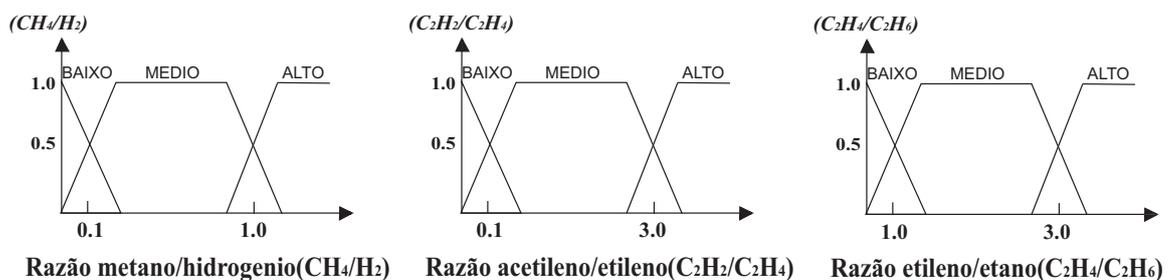
corretamente esse tipo de transição, especialmente quando estão ocorrendo mais de um tipo de falha (LIMA, 2005).

Dessa forma, a lógica fuzzy permite que as fronteiras existentes entre os diagnósticos possíveis para cada falha seja redefinida mediante cada aplicação, utilizando o conceito de grau de pertinência, sendo os sistemas fuzzy, alvo de muitos estudos voltados para o diagnóstico de falta em transformadores (DUKARM, 1993; ISLAM, WU e LEDWICH, 2000; MORAIS, ROLIM e SILVA, 2005; NETO e CHIGANER, 1999; WANG, LIU e GRIFFIN, 1998; XU et al., 1997; HUANG, YANG e HUANG, 1997; AMORA et al., 2009; MIRANDA e CASTRO, 2005; HOOSHMAND et al., 2012).

Usando a teoria de inferência fuzzy, pode ser construído um conjunto fuzzy baseado no diagnóstico de faltas. O foco principal é sintonizar as funções de pertinência baseadas nos métodos e experiências da análise de gases dissolvidos no óleo. Convencionalmente isto é feito manualmente (DUKARM, 93), e depois automaticamente utilizando métodos matemáticos sofisticados como computação evolucionária (evolutionary computing) e algoritmos genéticos (ARANTES, 2005).

Dukarm verificou que uma abordagem utilizando técnicas fuzzy e sistemas de inferência fuzzy se aplicariam aos métodos de análise de gases dissolvidos já apresentados. Como os métodos de diagnóstico de Rogers e Dornenburg baseiam-se em valores limiares, para que se possa aplicar a lógica fuzzy juntamente com estes métodos, as razões de gases são aplicadas a um sistema de inferência (vide Figuras 15 e 16) no qual as funções de pertinência associadas aos conjuntos fuzzy são relacionadas às variáveis consideradas. Para cada gás é associado um conjunto denominado Baixo, Médio ou Alto de acordo com seu nível de concentração no óleo (ARANTES, 2005).

Figura 15 – Diagramas fuzzy para as razões de Rogers



Fonte – Adaptado de Arantes, 2005

Figura 16 – Sistema de inferência para diagnóstico de falhas em transformadores para Rogers



Fonte – Adaptado de Arantes, 2005

### 2.5.3 Árvores de decisão

Um campo vasto e promissor, porém ainda pouco explorado é a utilização de algoritmos de mineração de dados para o diagnóstico de falhas incipientes em transformadores de força. Nesse contexto destaca-se a utilização de árvores de decisão.

Mineração de dados (DM, do inglês *Data Mining*) é a ciência associada à exploração de conjuntos grandes e complexos de dados, a fim de descobrir padrões úteis (ROKACH e MAIMON, 2008). Pesquisadores estão continuamente à procura de melhores técnicas para tornar o processo de data mining mais eficiente e preciso. Uma das abordagens mais promissoras e mais populares é o uso de DT. As árvores de decisão são técnicas simples, mas bem-sucedidas para prever e explicar a relação entre algumas medições sobre um item e seu valor-alvo. Além de seu uso em mineração, uma DT é uma ferramenta altamente eficaz em outras áreas, como a identificação de texto, extração de informações, aprendizagem de máquina e reconhecimento de padrões (AMORA, 2013).

Vários métodos de data mining como, por exemplo, uma DT são denominados métodos supervisionados, pois tentam descobrir a relação entre os atributos de entrada (às vezes chamados de variáveis independentes) e um atributo de destino (por vezes referido como uma variável dependente). A relação que é descoberta está representada na estrutura referida como um modelo. Usualmente os modelos podem descrever e explicar fenômenos que estão escondidos no conjunto de dados, e que pode ser usado para prever o valor do atributo alvo quando os valores dos atributos de entrada são conhecidos (ROKACH E MAIMON, 2008).

Os métodos não supervisionados tentam realizar a descoberta de relações entre os dados sem a utilização de variáveis dependentes, ou seja, sem a utilização de valores de saída ou rótulos pré-estabelecidos (KOHAVI e PROVOST, 1998).

Os modelos supervisionados podem ser divididos em modelos de classificação e modelos de regressão. Os modelos de regressão promovem o mapeamento do espaço de entrada em um domínio de valor real. Por exemplo, um modelo de regressão pode prever a procura (quantidade) de um determinado produto, dadas as suas características. Por outro lado, os classificadores promovem o mapeamento do espaço de entrada em classes prédefinidas. Por exemplo, os classificadores podem ser usados para classificar os credores de hipotecas como bons (hipoteca paga no tempo estipulado) e ruins (atraso de pagamento). Entre as muitas alternativas para representar classificadores, há, por exemplo, Máquinas de vetor de suporte, DT, função algébrica, e outros (ROKACH E MAIMON, 2008).

Em Amora (2013), foi proposta uma metodologia alternativa para a solução de problemas de classificação adotando DT, pela utilização, inicialmente, de um processo simplificado de aumento de dimensão no modelo do sistema e, posteriormente, o uso da capacidade de seleção de atributos durante a construção da própria DT, para então serem extraídas regras de classificação.

O novo método propõe uma nova representação de um problema, utilizando novos atributos ou relações não lineares das variáveis originais do problema, e, portanto, aumentando a dimensão do sistema. Após isso, é utilizada uma DT que através de sua capacidade intrínseca de seleção escolherá as variáveis mais relevantes para a solução do problema. A própria topologia determinada para a DT construída, também, permite a extração direta de regras de classificação (AMORA, 2013).

No método desenvolvido, utilizando uma DT para o diagnóstico de equipamentos com isolamento a óleo como os transformadores, foi adotado um banco de dados com 162 amostras. Sendo 117 casos do conjunto de dados IEC TC 10 (DUVAL e PABLO, 2001). Também são utilizados 39 casos retirados das Tabelas de I a III da referência (DUVAL, 2002). O acerto obtido para a classificação das amostras acima foi de 98%.

Para os 117 casos do conjunto de dados IEC TC 10 (DUVAL e PABLO, 2001), o algoritmo obteve êxito na classificação de 116 amostras. Para os 39 casos retirados das Tabelas de I a III da referência (DUVAL, 2002), o algoritmo obteve êxito na classificação de 37 amostras e para os 6 casos restantes o êxito foi de 100%.

## 2.6 Considerações Parciais

Nesse capítulo foram apresentados tópicos relacionados a falhas incipientes em transformadores de força. Inicialmente abordou-se o óleo isolante, suas propriedades e tipos disponíveis para utilização em transformadores, apresentado ao final da explanação, as diferenças existentes entre os óleos naftênico e parafínico bem como os motivos que levam a degradação de ambos.

Posteriormente, foram levantadas as falhas elétricas e térmicas que podem acometer o transformador de força, bem como os métodos tradicionais utilizados pelas normas com a finalidade de prover o diagnóstico de falhas incipientes e suas limitações para aplicação.

Na seção 2.5 foram apresentadas algumas das novas metodologias propostas para otimizar o processo de diagnóstico de falhas incipientes em transformadores de força, com a utilização de redes neurais, sistemas de inferência fuzzy e árvores de decisão. Dentre as novas metodologias disponíveis, destacou-se a proposta em (AMORA, 2013), como sendo uma metodologia bastante promissora e pouco explorada.

Neste trabalho, buscou-se o aprimoramento da técnica desenvolvida por AMORA (2013) com a utilização de árvore de decisão para classificação DGA, sendo que primeiramente obteve-se um aumento na dimensionalidade dos dados de entrada, conforme o teorema de Cover e posterior indução da árvore com base no novo espaço de entradas.

A ideia do aprimoramento da técnica desenvolvida por Amora (2013), está no fato de que por se tratar de um algoritmo de mineração de dados, árvores de decisão, recorrem ao máximo de informações contidas nos dados disponíveis como entrada para a melhor partição dos conjuntos de saídas.

Durante o Capítulo 3, são apresentadas considerações acerca de algoritmos de mineração de dados, árvores de decisão e das implementações de árvores utilizando os algoritmos ID3 e o C4.5. No Capítulo 4, será apresentado a forma como se desenvolveu a indução da árvore de decisão proposta utilizando o algoritmo C4.5 implementado no software SIPINA.

No Capítulo 5 são discutidos os resultados da implementação da árvore proposta, bem como uma comparação quanti-qualitativa com demais métodos de análise DGA.

### 3 DATA MINING E ÁRVORES DE DECISÃO

O uso de árvores de decisão como ferramenta para mineração de dados é uma prática crescente e promissora dentro do processo de descoberta do conhecimento em base de dados (KDD, do inglês *Knowledge Discovery in Databases*). Estudos indicam o uso promissor da ferramenta nos ramos da saúde, otimização de custos e tendências. Em Amora (2013), foi desenvolvida uma árvore de decisão para a classificação de falhas incipientes em transformadores com resultado satisfatório, fato este, motivador para o aprofundamento da pesquisa e formatação do presente trabalho.

Na seção 3.1 são apresentadas as etapas do processo de Descoberta de Conhecimento contido em bases de dados. Para encontrar respostas ou extrair conhecimento, existem diversos métodos de Data Mining disponíveis na literatura. Classificação, Modelos de Relacionamento entre Variáveis, Análise de Agrupamento, Sumarização, Modelo de Dependência, Regras de Associação e Análise de Séries Temporais são alguns destes.

Na seção 3.2 serão vistos os principais conceitos sobre árvores de decisão, como são formadas, os processos de indução, e os métodos de poda utilizados para redução do tamanho da árvore formada.

Durante a seção 3.3 serão apresentados dois dos principais algoritmos utilizados para indução de árvores de decisão, o ID3 e o C4.5, que é uma melhoria do anterior. Por fim, na seção 3.4 serão apresentadas as considerações parciais acerca do capítulo 3.

#### 3.1 Descoberta de Conhecimento em Base de Dados

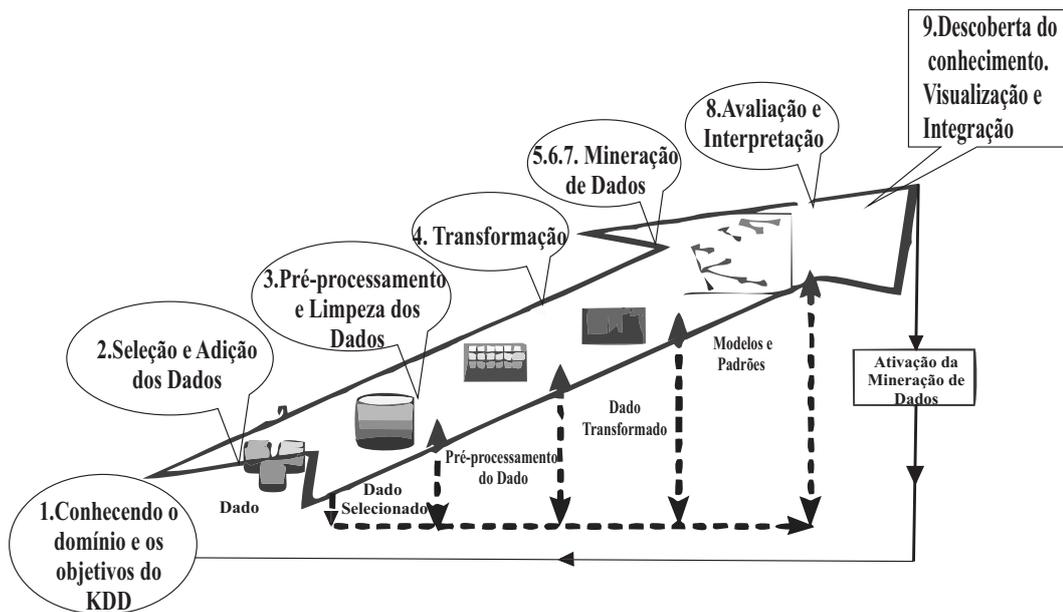
A Descoberta de Conhecimento em Base de Dados foi definida como um processo não trivial de extração de informações implícitas, previamente desconhecidas e potencialmente útil, a partir dos dados armazenados em um banco de dados (FAYYAD et al. 1996).

O processo KDD é constituído de várias etapas, que são executadas de forma interativa e iterativa. As etapas são interativas porque envolvem a cooperação da pessoa responsável pela análise de dados, cujo conhecimento sobre o domínio orientará a execução do processo. Por sua vez, a iteração deve-se ao fato de que, com frequência, esse processo não é executado de forma seqüencial, mas envolve repetidas seleções de parâmetros e conjunto de dados, aplicações das técnicas de Data Mining e posterior análise dos resultados obtidos, a fim de refinar os conhecimentos extraídos (BRANCHMAN E ANAND, 1996).

A mineração de dados é considerada como a etapa central para o processo de KDD. Dessa forma, o termo mineração de dados (Data Mining) pode ser considerado como sinônimo do processo completo.

A Figura 17 ilustra as etapas do processo de Descoberta de Conhecimento em Base de Dados conforme (FAYYAD et al., 1996b)

Figura 17 – Descoberta do Conhecimento em Base de Dados segundo(FAYYAD et al., 1996b)



Fonte – Adaptada de ROKACH E MAIMON, 2014

O processo de KDD é estruturado em nove etapas, que vão desde a compreensão do seu espaço de busca, com a elaboração dos objetivos do processo até a etapa final em que se tem a descoberta do conhecimento.

A Tabela 6, apresenta de forma resumida as atribuições de cada uma das etapas do processo de KDD apresentadas na Figura 20.

### 3.1.1 Data Mining

Data Mining, ou Mineração de Dados, pode ser entendido como o processo de extração de informações, sem conhecimento prévio, de um grande banco de dados e seu uso para tomada de decisões. É uma metodologia aplicada em diversas áreas que usam o conhecimento, como empresas, indústrias e instituições de pesquisa. Data Mining define o processo automati-

Tabela 6 – Etapas do Processo de Descoberta do Conhecimento

<b>Etapa</b>	<b>Nome</b>	<b>Atribuição</b>
1	Entender o domínio do conhecimento e identificando o objetivo do KDD	Nessa etapa, deve-se verificar o que se quer extrair e o que se planeja concluir ao final. Quanto mais definido isso estiver, maior será o filtro a ser utilizado na base de dados e mais detalhado será o resultado, como em um processo de análise estatística comum.
2	Seleção dos dados	É a primeira no processo de descobrimento de informação. Nesta fase é escolhido o conjunto de dados, pertencente a um domínio, contendo todos os possíveis atributos e registros que farão parte da análise. O processo de seleção é bastante complexo, uma vez que os dados podem vir de uma série de fontes diferentes e podem possuir os mais diversos formatos.
3	Pré-processamento e limpeza dos dados	Nesta etapa deverão ser realizadas tarefas que eliminem dados redundantes e inconsistentes, recuperem dados incompletos e avaliem possíveis dados discrepantes ao conjunto (outliers). Nesta fase também são utilizados métodos de redução ou transformação para diminuir o número de variáveis envolvidas no processo, visando com isto melhorar o desempenho do algoritmo de análise.
4	Transformação dos Dados	Após serem selecionados, limpos e pré-processados os dados necessitam ser armazenados e formatados adequadamente para que os algoritmos de aprendizado possam ser aplicados. Os dados que estão dispersos devem ser agrupados em um repositório único
5,6 e 7	Mineração de dados	É a etapa de maior importância no processo de KDD. Data mining é a exploração e análise, de forma automática ou semi-automática, de grandes bases de dados com objetivo de descobrir padrões e regras. Dentre as técnicas disponíveis para serem utilizadas como Data mining, destaca-se: (i) Árvores de decisão, (ii) Regras de indução, (iii) Redes Neurais, (iv) Análise de regressão e (v) Análise de agrupamentos.
8	Avaliação	Nessa etapa são avaliados e interpretados os padrões extraídos, em relação aos objetivos definidos primariamente. Essa etapa centra-se na compreensão e busca da utilidade do modelo induzido.
9	Usando o conhecimento adquirido	É a etapa de aplicação do conhecimento adquirido em um processo novo ou pré-existente. O sucesso desta etapa determina a eficácia de todo o processo.

zado de captura e análise de grandes conjuntos de dados para extrair um significado. (SFERRA E CORRÊA, 2003)

A maioria das técnicas de mineração de dados baseia-se na aprendizagem indutiva na qual um modelo é construído explicitamente ou implicitamente, generalizado a partir de um número suficiente de exemplos de treinamento (ROKACH E MAIMON, 2014).

Para extração do conhecimento, existem diversos métodos de Data Mining disponíveis na literatura. A maioria desses métodos é baseada em técnicas de aprendizado de máquina, reconhecimento de padrões e estatística. Essas técnicas vão desde a tradicional estatística multivariada, como análise de agrupamentos e regressões, até modelos mais atuais de aprendizagem, como redes neurais, lógica difusa e algoritmos genéticos (SFERRA E CORRÊA, 2003).

É útil distinguir entre dois tipos principais de mineração de dados: (i) orientado para verificação (o sistema verifica a hipótese do usuário) e (ii) orientado a descoberta, ou seja, o sistema encontra novas regras e padrões de forma autônoma (ROKACH E MAIMON, 2014).

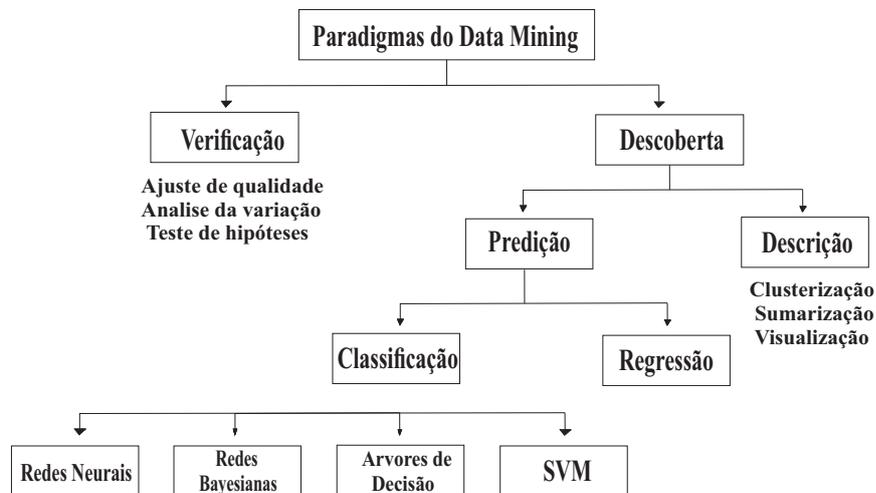
A Figura 18 ilustra essa taxonomia. Cada tipo tem sua própria metodologia. Os métodos de descoberta, que identificam automaticamente padrões nos dados, envolvem métodos de predição e descrição. Métodos de descrição concentram-se em entender a forma como os dados subjacentes funcionam enquanto os métodos orientados para a previsão visam construir um modelo comportamental para obter amostras novas e não vistas afim de prever valores de uma ou mais variáveis relacionado à amostra. Alguns métodos orientados para a predição, no entanto, também podem contribuir para a compreensão dos dados (ROKACH E MAIMON, 2014).

Embora a maioria das técnicas orientadas a descobertas utilizem a aprendizagem indutiva como discutido acima, os métodos de verificação avaliam uma hipótese proposta por uma fonte externa, como especialista (ROKACH E MAIMON, 2014).

A maioria dos problemas de mineração de dados está preocupada com a seleção de uma hipótese em vez de testar um conhecido. Enquanto um dos principais objetivos da mineração de dados é a identificação do modelo, métodos estatísticos geralmente se concentram na estimação do modelo (ELDER E PREGIBON, 1996).

Em mineração de dados, uma árvore de decisão é um modelo preditivo que pode ser usado para representar tanto classificadores quanto modelos de regressão. As árvores de decisão são modelos hierárquicos de decisões e suas consequências. Quando uma árvore é usada para tarefas de classificação, é mais comumente referida como árvore de classificação. Quando é usado para tarefas de regressão, é chamada de árvore de regressão (ROKACH E MAIMON,

Figura 18 – Taxonomia do método de Data Mining



Fonte – Adaptado de ROKACH E MAIMON, 2014

2014).

A árvore de decisão é útil como técnica exploratória, no entanto, não tenta substituir métodos estatísticos tradicionais existentes. Existem muitas outras técnicas que podem ser utilizadas para classificar ou prever a afiliação de instâncias com um conjunto predefinido de classes, como redes neurais artificiais ou máquinas de vetor de suporte (ROKACH E MAIMON, 2014).

Na seção seguinte, serão apresentados os conceitos acerca de Árvores de Decisão, visto que esta será a ferramenta de mineração de dados utilizada no presente trabalho para classificação das falhas internas em transformadores de força a partir dos gases dissolvidos em amostras de óleo do equipamento.

### 3.2 Árvore de Decisão

Uma Árvore de decisão representa um sistema de decisão multiestágios no qual as classes são sequencialmente rejeitadas até ser alcançada uma classe final de aceitação, durante a apresentação de uma instância (vetor de atributos) a ser classificada. No final, o espaço de entrada é dividido em regiões distintas, correspondendo às classes, de maneira sequencial. Durante a apresentação de um vetor para classificação, a pesquisa da região a ser associada a um parâmetro do vetor é obtida através de uma sequência de decisões ao longo de um caminho de nós, numa árvore apropriadamente construída (THEODORIDIS e KOUTROUMBAS, 2009).

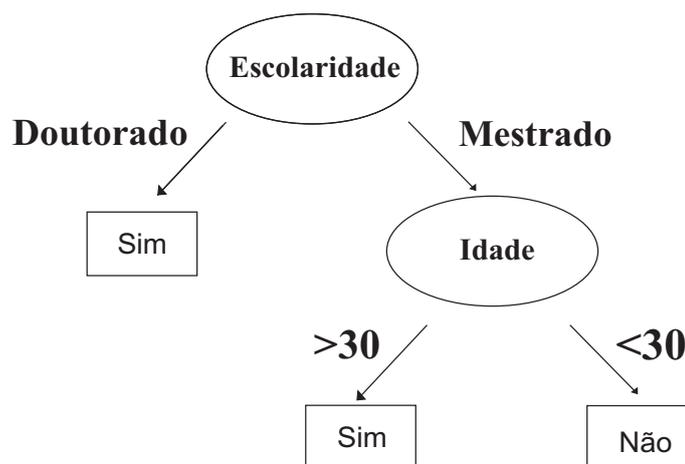
Uma árvore de decisão é um classificador expresso como uma partição recursiva do espaço de instância. A árvore de decisão consiste em nós que formam uma árvore roteada, ou seja, é uma árvore direcionada por um nó chamado "raiz" que não recebe bordas. Todos os outros nós têm exatamente uma borda de entrada. Um nó com bordas de saída é referido como um nó "interno" ou um nó de "teste". Todos os outros nós são chamados de "folhas" (também conhecidas como "terminal") ou (nós de "decisão").

Em uma árvore de decisão, cada nó interno divide o espaço em dois ou mais sub-espacos de acordo com uma determinada função discreta dos valores de atributos de entrada. No caso mais simples e mais freqüente, cada nó de "teste" considera um único atributo, de modo que o espaço da instância seja particionado de acordo com o valor dos atributos. No caso dos atributos numéricos, a condição refere-se a um intervalo. (ROKACH E MAIMON, 2014)

Cada folha é atribuída a uma classe que representa o valor mais adequado. Alternativamente, a folha pode conter um vetor de probabilidade (vetor de afinidade) indicando a probabilidade de o atributo alvo ter um certo valor.

A Figura 19 descreve um exemplo de uma árvore de decisão que prevê se uma pessoa é rica ou não com base em sua escolaridade. Os nós internos são representados como círculos, enquanto as folhas são denotadas como triângulos. Dois ou mais ramos podem surgir de cada nó interno.

Figura 19 – Exemplo de Árvore de Decisão



Fonte – ROKACH E MAIMON, 2014

Cada nó corresponde a uma determinada característica e os ramos correspondem a

uma gama de valores. Esses intervalos de valores devem ser mutuamente exclusivo e completo. Essas duas propriedades, disjunção e integridade são importantes, uma vez que garantem que cada instância de dados seja mapeada para uma classificação.

As instâncias são classificadas navegando-as a partir da raiz da árvore até uma folha de acordo com o resultado dos testes ao longo do caminho. Começando pela raiz de uma árvore, testa-se a característica que corresponde ao referido nó e define-se qual será a ramificação a ser seguida com base no valor observado para a característica correspondente.

Após o primeiro teste, considera-se o nó em que o dado aparece e repetimos as mesmas operações para este nó, até chegarmos a uma folha. Cada nó é rotulado com o atributo que testa, e seus ramos são rotulados com seus valores correspondentes.

Em caso de atributos numéricos, as árvores de decisão podem ser geometricamente interpretado como uma coleção de hiperplanos, cada um ortogonal a um dos eixos (ROKACH E MAIMON, 2014).

Uma DT típica separa o espaço de entradas em hiperplanos com retas paralelas aos eixos, como exemplificado na Figura 20. A sequência de decisões é aplicada para cada atributo da instância apresentada à árvore, com os testes de decisão associados aos nós sendo na forma (AMORA,2013):

Se  $a_i \theta \alpha$  então  $c_1$ , senão  $c_2$

onde:

$a_i$  representa o atributo avaliado;

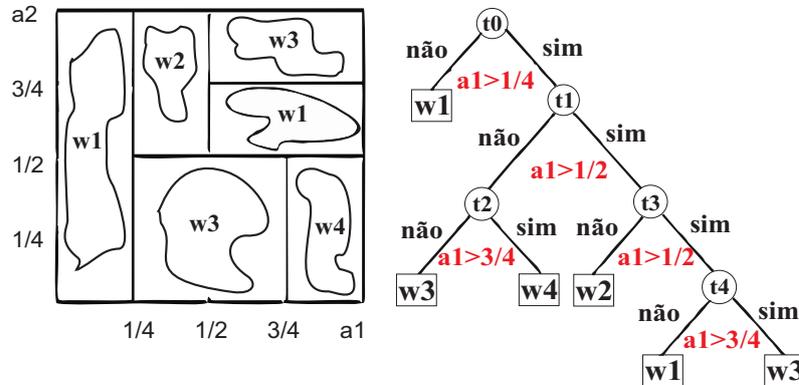
$\theta$  é a operação lógica testada ( $=, \neq, \leq, \geq, <, >$ );

$\alpha$  é um valor limite; e

$c_1$  e  $c_2$  representam “caminhos” distintos na árvore que levam a outros nós na árvore que podem representar um outro nó de teste ou então um nó de “folha” que representa uma classe.

Uma árvore com nós descritos pela relação acima é denominada de Árvore de Classificação Binária (ACB), pois existem apenas dois caminhos que podem derivar de cada nó de teste. Na Figura 20, além da partição do espaço, também é apresentada a ACB obtida, indicando as regras de classificação (AMORA, 2013).

Figura 20 – Partição do espaço de variáveis e regras obtidas de uma árvore de classificação binária



Fonte – Adaptado de AMORA, 2013

### 3.2.1 Indução de Árvores de Decisão

Várias aplicações com a utilização de Inteligência Artificial para tarefas de importância prática são baseadas na construção de um modelo de conhecimento que é utilizado por um especialista humano.

O processo de indução de árvores consiste na utilização de algoritmos para a concepção de um modelo de decisão matemático, baseado em questionamentos nó-a-nó, desde a raiz até a folha da árvore com a finalidade de classificação correta das amostras apresentadas.

Um algoritmo de indução, é uma entidade que obtém um conjunto de treinamento e forma um modelo que generaliza a relação entre os atributos de entrada e a saída desejada. Normalmente, o objetivo é encontrar a melhor árvore de decisão, minimizando o erro de generalização. No entanto, as funções também podem ser definidas, por exemplo, para minimizar o número de nós ou minimizar a profundidade média da árvore. (ROKACH E MAIMON, 2014)

Existem duas formas de indução de uma DT: (i) top-down, onde a árvore é construída de cima para baixo na busca do melhor atributo e (ii) bottom-up, com construção inversa, ou seja de baixo para cima, sendo dado preferência para a primeira forma na maioria dos algoritmos de indução.

Existem vários algoritmos de indução de árvores de decisão do tipo Top-down, como o ID3 (Quinlan, 1986), C4.5 (Quinlan, 1993), CART (Breiman et al. 1984). Alguns algoritmos de indução, consistem em duas fases conceituais: Crescimento e podas, que é o caso do C4.5 e CART. Outros, executam apenas a fase de crescimento.

A Figura 21 ilustra um pseudo-código típico para uma indução do tipo top-down de um algoritmo de uma árvore de decisão usando os processos de crescimento e poda, que serão apresentados em seções posteriores. Note que estes algoritmos são "gulosos", ou seja, procuram localmente o melhor atributo para proferir a divisão, independente das divisões posteriores que poderão ocorrer durante o processo de indução da árvore, construindo assim uma árvore recursiva (também conhecida como divisão e conquista). Em cada iteração, o algoritmo considera a partição do conjunto de treinamento usando o resultado de atributos discretos de entrada.

A seleção dos atributos mais apropriados é feita de acordo com algumas medidas de divisão. Após a seleção de uma divisão apropriada, cada nó subdivide o conjunto de treinamento em subconjuntos menores, até que um critério de parada seja satisfeito. (ROKACH E MAIMON, 2014)

Uma alternativa para limitar o tamanho de um DT no seu processo de indução é aplicar um método de poda após a construção da árvore, de forma a reduzir o tamanho desta, eliminando nós e ramos com menor impacto na precisão de classificação realizada pela árvore. A ideia de poda foi originalmente proposta por (BREIMAN et al., 1984).

Estudos demonstram que a poda permite a melhoria na capacidade de generalização da árvore, mesmo em situações de ruído nos dados. No capítulo 4, serão avaliados os processos de formação de árvores de indução com e sem a ocorrência de poda, e as implicações decorrentes da não utilização deste processo.

O algoritmo de indução segue na formação da árvore de decisão até que um dos critérios de parada abaixo seja atendido:

- Todas as instâncias no conjunto de treinamento pertencem a um único valor de classe;
- O tamanho máximo da árvore foi alcançado;
- O número de casos no nó terminal é inferior ao mínimo número de casos para nós pai;
- Se o nó foi dividido, o número de casos em um ou mais nós filhos seria menor que o número mínimo de casos para nós infantis;
- O melhor critério de divisão não é superior a um determinado limite.

Figura 21 – Exemplo de pseudocódigo para construção de Árvore de Decisão

```

Procedimento ArvoreIndutor( $X,A,y$ )
     $T=ArvoreCrescimento(X,A,y)$ 
    Return  $ArvoreCrescimento(X,A,y)$ 

Procedimento ArvoreCrescimento( $X,A,y$ )
    Criar uma árvore  $T$ 
    Se um dos critérios de parada é alcançado Então
        Faça o nó raiz em  $T$  como um nó folha associado ao valor mais comum
        de  $y$  em  $X$ , como classe.
    Senão
        Encontre uma função discreta  $f(A)$  dos valores dos atributos de
        entrada que divida  $X$  conforme uma métrica de divisão adotada.
        Se métrica de divisão  $\geq$  Limiar Então
            Rótulo do nó raiz em  $T$  é  $f(A)$  .
            Para cada resultado  $v_i$  de  $f(A)$  faça
                 $Sub-arvore_i = ArvoreCrescimento(f(A)=v_i, X,A,y)$ .
                Conecte o nó raiz de  $T$  a Subarvore, com uma aresta
                que é denominada como  $vi$ .
            Fim do para
        Senão
            Marque o nó raiz em  $T$  como uma folha com o valor mais
            comum de  $y$  em  $X$  como a classe.
        Fim do Se
    Fim do Se
    Retorna  $T$ 

Procedimento ArvorePoda( $X,T,y$ )
    Repita até
        Selecione um nó  $t$  em  $T$  onde a poda é máxima segundo o critério de
        evolução.
        Se  $t \neq 0$  Então
             $T=Poda(T,t)$ 
        Fim do Se
    Até t=0
    Retorna  $T$ 

```

### 3.2.2 Critérios utilizados para divisão de Árvores de Decisão

Na maioria dos algoritmos de indução para árvores de decisão, as funções discretas de divisão são univariadas, ou seja, um nó interno é dividido de acordo com o valor de um único atributo (ROKACH E MAIMON, 2014). Conseqüentemente, o algoritmo busca o melhor atributo sobre o qual executará a divisão. Existem alguns critérios para a realização da busca do atributo, que podem ser caracterizados:

- de acordo com a origem da medida: Teoria da informação, dependência, e distância;
- de acordo com a estrutura da medida: Critérios baseados em impurezas, normalização baseada na impureza e critérios binários.

#### 3.2.2.1 Critérios baseados em impurezas

Dada uma variável aleatória  $x$  com  $k$  valores discretos, distribuídos de acordo com  $P = (p_1, p_2, \dots, p_k)$ , uma medida de impureza é uma função  $\Phi : [0, 1]^k \rightarrow R$ , tal que satisfaça as seguintes condições:

- $\Phi(P) \geq 0$ ;
- $\Phi(P)$  é mínimo se  $\exists i$  quando  $p_i = 1$ ;
- $\Phi(P)$  é máximo se  $\forall i, 1 \leq i \leq k, p_i = 1/k$ ;
- $\Phi(P)$  é simétrico em relação aos componentes de  $P$ ;
- $\Phi(P)$  é diferenciável em todo seu domínio;

Cabe ressaltar que se o vetor de probabilidade tiver um componente cuja a variável  $x$  obtém apenas um valor, então a variável é definida como pura. Por outro lado, se todos os componentes forem iguais, o nível de impureza atingirá o valor máximo.

Dado um conjunto de treinamento  $S$  o vetor de probabilidade do atributo alvo  $y$  é definido como em (ROKACH E MAIMON, 2014) por:

$$P_y(S) = \frac{|\sigma_{y=c_1}S|}{|S|}, \dots, \frac{|\sigma_{y=c_{|dom(y)|}}S|}{|S|} \quad (3.1)$$

A melhor divisão para o atributo discreto  $a_i$  é definida como sendo a redução na

impureza do atributo alvo depois do particionamento  $S$  de acordo com os valores  $v_{i,j} \in dom_{ai}$ :

$$\Delta\Phi(ai, S) = \Phi(P_y(S)) - \sum_{j=1}^{|dom_{ai}|} \frac{|\sigma_{ai=v_{i,j}}S|}{|S|} \cdot \Phi(P_y(\sigma_{ai=v_{i,j}}S)) \quad (3.2)$$

### 3.2.2.2 Ganho de Informação

Ganho de informação é um critério para medida que usa a entropia (proveniente da teoria da informação) como medida de impureza.

$$ganho(ai, S) = Entropia(y, S) - \sum_{v_{i,j} \in dom_{ai}} \frac{|\sigma_{ai=v_{i,j}}S|}{|S|} \cdot Entropia(y, \sigma_{ai=v_{i,j}}S) \quad (3.3)$$

sendo a Entropia dada por:

$$Entropia(y, S) = \sum_{c_j \in dom_{(y)}} - \frac{|\sigma_{y=c_j}S|}{|S|} \cdot \log_2 \frac{|\sigma_{y=c_j}S|}{|S|} \quad (3.4)$$

O ganho de informação está intimamente relacionado com a estimativa de máxima verossimilhança, que é um método estatístico popular usado para fazer inferências sobre os parâmetros da distribuição de probabilidade subjacente de uma amostra de um conjunto de dados (ROKACH E MAIMON, 2014).

### 3.2.2.3 Índice Gini

O índice Gini é um critério baseado em impurezas que mede as divergências entre as distribuições de probabilidade dos valores em relação aos atributos alvo (ROKACH E MAIMON, 2014).

$$Gini(y, S) = 1 - \sum_{c_j \in dom_{(y)}} \left( \frac{|\sigma_{y=c_j}S|}{|S|} \right)^2 \quad (3.5)$$

Logo, o critério para seleção do atributo  $ai$  será:

$$GanhoGini(ai, S) = Gini(y, S) - \sum_{v_{i,j} \in dom_{ai}} \frac{|\sigma_{ai=v_{i,j}}S|}{|S|} \cdot Gini(y, \sigma_{ai=v_{i,j}}S) \quad (3.6)$$

### 3.2.2.4 Razão do Ganho

Os critérios baseados em impurezas, apresentados anteriormente, é tendencioso para atributos com valores de domínio maiores, ou seja, prefere atributos de entrada com muitos valores em detrimento aqueles com menos valores (QUINLAN, 1986).

Por esse motivo, faz-se necessário a normalização das medidas baseadas em impurezas. A relação de ganho normaliza o ganho de informação da seguinte forma (QUINLAN, 1993):

$$RazoGanho(ai, S) = \frac{GanhodaInformao(ai, S)}{Entropia(ai, S)} \quad (3.7)$$

Observe que essa relação não é definida quando o denominador é zero. Além disso, a proporção pode tender a favorecer atributos para os quais o denominador é muito pequeno. Consequentemente, sugere-se que a relação seja realizada em duas etapas. Primeiro, o ganho de informação é calculado para todos os atributos. Como uma consequência, deve-se considerar apenas os atributos que obtiveram um valor satisfatório do ganho médio de informação. Dessa forma, o atributo que obteve a melhor razão do ganho é selecionado. Quinlan (1988) mostrou que a relação de ganho tende a superar os critérios simples de obtenção de informações, tanto na precisão quanto em termos de complexidade classificadora. Uma penalidade é avaliada para a informação ganho de um atributo contínuo com muitas divisões possíveis.

### 3.2.2.5 Critério Ortogonal

O critério ortogonal foi apresentado por (FAYYAD E IRANI, 1992). Este critério binário é definido como:

$$ORT(ai, dom_1(ai), dom_2(ai)) = 1 - \cos \theta(P_{y,1}, P_{y,2}) \quad (3.8)$$

sendo  $\theta (P_{y,1} \text{ e } P_{y,2})$  o ângulo entre dois vetores  $P_{y,1}$  e  $P_{y,2}$ . Estes vetores representam a distribuição de probabilidade do atributo de destino na partição  $\sigma_{ai \in dom_1(a1)}$  e  $\sigma_{ai \in dom_2(a1)}$  respectivamente.

### 3.2.3 Poda em Árvores de Decisão

Empregar critérios de parada rigorosos, tendem a criar árvores de decisão pequenas e incompletas. Por outro lado, o uso de critérios de parada menos rigorosos tendem a gerar grandes árvores de decisão que podem gerar o *overfitting* no conjunto de treinamento. Para resolver este dilema, Breiman et al. (1984) desenvolveu uma metodologia de poda com base em um critério de parada menos rigorosos e permitindo que a árvore de decisão provocasse o *overfitting* no conjunto de treinamento. Então, a árvore formada é cortada de volta se tornando uma árvore menor, removendo sub-ramos que não contribuem para a generalização (ROKACH E MAIMON, 2014).

A opção por utilizar a poda em árvores de decisão, contempla também a otimização de recursos disponíveis, sem que necessariamente se perca a capacidade de generalização, impactando em perda de rendimento na atividade de classificação.

Existem várias formas de realizar poda em uma árvore de decisão, e todas elas são classificadas como pré-poda ou pós-poda. O método pré-poda é realizado durante o processo de construção da árvore, em que o processo pode simplesmente parar de dividir o conjunto de elementos e transformar o nó corrente em um nó folha da árvore. O ganho de informação, por exemplo, pode ser utilizado como critério de poda. Caso todas as divisões possíveis utilizando um atributo *A* gerem ganhos menores que um valor pré-estabelecido, então esse nó vira folha, representando a classe mais frequente no conjunto de exemplos (BASGALUPP, 2010).

A dificuldade é encontrar um valor adequado para a divisão, visto que um valor muito alto pode gerar uma árvore super-simplificada, enquanto que um valor muito baixo pode simplificar bem pouco a árvore (BASGALUPP, 2010).

Já o pós-poda é realizado após a construção da árvore de decisão, removendo ramos completos, onde tudo que está abaixo de um nó interno é excluído e esse nó é transformado em folha, representando a classe mais frequente no ramo. Para cada nó interno da árvore, o algoritmo calcula a taxa de erro caso a sub-árvore abaixo desse nó seja podada. Em seguida, é calculada a taxa de erro caso não haja a poda. Se a diferença entre essas duas taxas de erro for menor que um valor pré-estabelecido, a árvore é podada. Caso contrário, não ocorre a poda (FERREIRA, 2008).

Esse processo se repete progressivamente, gerando um conjunto de árvores podadas. Por fim, para cada uma delas é calculada a precisão na classificação de um conjunto de dados independente dos dados de treinamento (por exemplo, o conjunto de validação), e a árvore que

obtiver a melhor precisão será a escolhida (FERREIRA, 2008).

Nas sub-seções seguintes, serão abordadas duas das principais técnicas desenvolvidas para a realização de podas em árvores induzidas, que é a técnica de poda por Custo de Complexidade e o Erro Mínimo de Poda.

### 3.2.3.1 Poda por Custo de Complexidade

O modelo proposto em [BREIMAN et al.(1984)] é organizado em duas etapas. Na primeira etapa do algoritmo, é criada uma seqüência de árvores  $T_0, T_1, \dots, T_k$ , tomando por base os dados de treinamento, sendo que a árvore  $T_0$  é a árvore primariamente formada, antes da ocorrência da poda e  $T_k$  é a árvore raiz. Na segunda etapa, uma das árvores criadas é escolhida para ser a árvore com poda, baseado na estimativa do erro de generalização de cada uma.

A árvore  $T_{i+1}$  é obtida substituindo uma ou mais sub-árvores na árvore antecessora  $T_i$  com folhas adequadas. As sub-árvores que são podadas são aquelas que obtêm o menor aumento na taxa de erro aparente por folha podada, que é dado por:

$$\alpha = \frac{\varepsilon(\text{poda}(T, t), S) - \varepsilon(T, S)}{|\text{folhas}(T)| - |\text{folhas}(\text{poda}(T, t))|} \quad (3.9)$$

sendo  $\varepsilon(\text{poda}(T, t), S)$  a taxa de erro da árvore  $T$  sobre a amostra  $S$ ,  $\text{folhas}(T)$  indica o número de folhas em  $T$  e  $\text{poda}(T, t)$  denota a árvore obtida substituindo o nó  $t$  em  $T$  por uma folha adequada.

Na segunda fase, o erro de generalização de cada árvore podada é calculado. A melhor árvore podada é então selecionada. Se o conjunto de dados for grande o suficiente, é recomendado separar um conjunto para o treinamento e uma para poda. As árvores são construídas usando o conjunto de treinamento e avaliadas no conjunto de poda. Por outro lado, se o determinado conjunto de dados não for suficientemente grande, pode-se usar a metodologia de validação cruzada (ROKACH E MAIMON, 2014).

### 3.2.3.2 Erro mínimo de Poda

O método do Erro Mínimo de Poda (MEP), proposto por Niblett e Bratko (1986), envolve uma busca *bottom-up* dos nós internos. Esta técnica compara, em cada nó, a estimativa da taxa de erro de *l-probabilidades* da taxa de estimativa de erro  $l$  com e sem poda. A estimativa da taxa de erro de *l-probabilidades* é uma correção para a estimativa de probabilidade simples,

usando frequências. Se  $S_t$  denota as instâncias que atingiram uma folha  $t$ , então a taxa de erro esperada nesta folha é (ROKACH E MAIMON, 2008):

$$\varepsilon'(t) = 1 - \max_{c_1 \in \text{dom}(y)} \frac{|\sigma_{y=c_1} S_t| + l \cdot \text{Pr}(y=c_1)}{|S_t| + l} \quad (3.10)$$

sendo  $\text{Pr}(y=c_1)$  a probabilidade de se obter o valor  $c_1$ , e  $l$  denota o peso dado a essa probabilidade.

A taxa de erro de um nó interno é a média ponderada do erro da taxa de seus ramos. O peso é determinado de acordo com a proporção de instâncias ao longo de cada ramo. O cálculo é executado recursivamente para as folhas. Se um nó interno for podado, torna-se uma folha e sua taxa de erro é calculado diretamente usando a última equação. Consequentemente, é possível comparar a taxa de erro antes e depois da poda de um determinado nó interno. Se a poda neste nó não aumenta a taxa de erro, esta deve ser aceita.

### 3.3 Algoritmos para Indução de Árvores de Decisão

Nesta seção serão abordados os principais pontos acerca de dois algoritmos utilizados para indução de árvores de decisão. Inicialmente, será observado o algoritmo ID3 proposto em (QUINLAN, 1986), como sendo o precursor para o estudo do algoritmo C4.5 (QUINLAN, 1993) que será o algoritmo utilizado ao longo deste trabalho para a indução de árvores de decisão para diagnóstico dos gases dissolvidos em transformadores de força.

Certo de que existem outros algoritmos que induzem árvores de decisão como CART (BREIMAN et al., 1984), CHAID (KASS, 1980), QUEST (LOH e SHIH, 1997), bem como algoritmos que não utilizam a estrutura do tipo *Top-Down Induction of Decision Trees (TDIDT)*, como NBtree (KOHAVI, 1996), ADTree (FREUND e MASON, 1999), LMT (LANDWEHR et al., 2005) e BFTree (SHI, 2007), optou-se pela utilização do algoritmo C4.5 devido a facilidade de interpretação dos conceitos agregados, interoperabilidade com o algoritmo base (ID3) e fácil implementação, utilizando ferramentas computacionais de fácil programação e boa aceitação acadêmica e comercial.

#### 3.3.1 Algoritmo ID3

O ID3 (QUINLAN, 1986) é o algoritmo pioneiro em indução de árvores de decisão. Ele é um algoritmo recursivo e baseado em busca gulosa, procurando, sobre um conjunto de

atributos, aqueles que “melhor” dividem os exemplos, gerando sub-árvores. A principal limitação do ID3 é que ele só lida com atributos categóricos não-ordinais, não sendo possível apresentar a ele conjuntos de dados com atributos contínuos, por exemplo. Nesse caso, os atributos contínuos devem ser previamente discretizados. Além dessa limitação, o ID3 também não apresenta nenhuma forma para tratar valores desconhecidos, ou seja, todos os exemplos do conjunto de treinamento devem ter valores conhecidos para todos os seus atributos. Na prática, os conjuntos de dados possuem muitos valores desconhecidos (BASGALUPP, 2010).

Logo, para se viabilizar a utilização do ID3, é necessário um trabalho direcionado e oneroso no pré-processamento dos dados e tratamento da informação.

A principal vantagem do ID3 é a sua simplicidade. Por esse motivo, esse algoritmo é frequentemente utilizado para fins de ensino, porém apresentando algumas desvantagens (ROKACH E MAIMON, 2008):

- Não garante uma solução ótima pois pode ficar preso em um mínimo local ótimo, visto que utiliza a estratégia gulosa;
- Pode causar *overfit* na formação da árvore com os dados de treinamento;
- Geralmente produz pequenas árvores, mas nem sempre produz as menores árvores possíveis;
- Foi projetado para atributos nominais. Portanto, dados contínuos podem ser usado somente depois de convertê-los em dados nominais.

### 3.3.1.1 Indução da árvore no ID3

A estrutura básica do ID3 é iterativa. Um subconjunto do conjunto de treinamento chamado janela é escolhido aleatoriamente e uma árvore de decisão formada a partir dele. Esta árvore corretamente classifica todos os objetos da janela escolhida. Todos os outros objetos no conjunto de treinamento são então classificados usando a árvore induzida. Se a árvore fornecer a resposta correta para todos esses objetos, então é correto para todo o conjunto de treinamento e o processo termina, caso contrário, uma seleção dos objetos incorretamente classificados é adicionado à janela e o processo continua (QUINLAN, 1986).

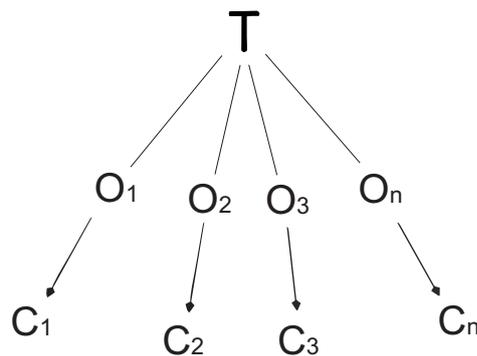
Desta forma, as árvores de decisão corretas foram encontradas após apenas algumas iterações para conjuntos de treinamento de até trinta mil objetos descritos em termos de até 50 atributos. Evidências empíricas sugerem que uma árvore de decisão correta geralmente é encontrada mais rapidamente por este método iterativo do que formando uma árvore diretamente

de todo o conjunto de treinamento. No entanto, observou que o quadro iterativo não pode ser garantido para convergir em uma árvore final a menos que a janela possa crescer para incluir a totalidade do conjunto de treinamento. Esta limitação potencial ainda não surgiu na prática (QUINLAN, 1986).

O grande problema para esse tipo de indução se dá em como formar uma árvore de decisão para uma coleção arbitrária  $C$  de objetos. Se  $C$  estiver vazio ou contiver apenas objetos de uma classe, a decisão mais simples seria reduzir a árvore apenas a uma folha marcada com a classe. Caso contrário, seja  $T$  qualquer teste em um objeto com possíveis resultados  $O_1, O_2, \dots, O_w$ , cada objeto em  $C$  dará um desses resultados para  $T$ , então  $T$  produz uma partição  $[C_1, C_2, \dots, C_w]$  de  $C$  com  $C_i$  contendo os objetos com resultado  $O_i$ .

Isso é representado graficamente pela forma de árvore da Figura 22. Se cada subconjunto  $C_i$  puder ser substituído por uma árvore de decisão para  $C_i$ , o resultado seria uma árvore de decisão para todos  $C$ . Além disso, enquanto dois ou mais  $C_i$  não forem vazios, cada  $C_i$  é menor que  $C$ . No pior caso, essa estratégia de divisão e conquista irá fornecer subconjuntos de um único objeto que satisfaçam o requisito de uma classe para uma folha. Portanto, desde que seja sempre possível encontrar um teste que ofereça uma partição não trivial de qualquer conjunto de objetos, esse procedimento sempre produzirá uma árvore de decisão que classifica corretamente cada objeto em  $C$  (QUINLAN, 1986).

Figura 22 – Estrutura de árvore para os objetos em  $C$



Fonte – QUINLAN, 1986

A escolha do teste é fundamental se a árvore de decisão for simples. Para o momento, um teste será restrito a ramificação dos valores de um atributo, de modo a escolher que o teste se resume a selecionar um atributo para a raiz da árvore. Para o ID3, adotou-se um método que

depende de dois pressupostos. Uma vez que C contem objetos  $p$  da classe "P" e  $n$  da classe "N", denota-se que:

- Qualquer árvore de decisão correta para C classificará objetos na mesma proporção que sua representação em C. Um objeto arbitrário será determinado a pertencer a classe P com probabilidade  $p / (p + n)$  e para classe N com probabilidade  $n / (p + n)$ ;
- Quando uma árvore de decisão é usada para classificar um objeto, ela retorna uma classe. Uma árvore de decisão pode assim ser considerada como fonte de uma mensagem 'P' ou 'N', com a informação necessária esperada para gerar esta mensagem dada por:

$$I(p, n) = -\frac{p}{p+n} \cdot \log_2 \frac{p}{p+n} - \frac{n}{p+n} \cdot \log_2 \frac{n}{p+n} \quad (3.11)$$

Se atributo A com valores  $[A_1, A_2, \dots, A_v]$  é usado para a raiz da árvore de decisão, irá dividir C em  $[C_1, C_2, \dots, C_v]$  onde  $C_i$  contém esses objetos em C que tem valor  $A_i$  de A. Para  $C_i$  que contém  $p_i$  objetos da classe P e  $n_i$  da classe N. A informação esperada necessária para a subárvore  $C_i$  é  $I(p_i, n_i)$ , é dada pela fórmula abaixo:

$$E(A) = \sum_{i=1}^v \frac{p_i + n_i}{p + n} I(p_i, n_i) \quad (3.12)$$

sendo que o peso para o  $i$ -ésimo ramo é a proporção dos objetos em C que pertencem para  $C_i$ . A informação obtida por ramificação em A é, portanto:

$$gain(A) = I(p, n) - E(A) \quad (3.13)$$

Em regra geral, é boa a escolha do atributo ligado ao ramo em que se ganha a maior informação. O ID3 examina todos os atributos candidatos e escolhe A para maximizar o ganho (A), forma a árvore e finalmente usa o mesmo processo de forma recursiva para formar árvores de decisão para os subconjuntos residuais  $C_1, C_2, \dots, C_w$ .

### 3.3.2 Algoritmo C4.5

O C4.5 é a evolução do ID3. Usa a razão de ganho como critério para a divisão dos nós da árvore. A divisão cessa quando o número dos casos a serem divididos estiver abaixo

de um determinado limite. A Poda baseada em erro é realizada após a fase de crescimento. O C4.5 pode lidar com atributos numéricos, bem como induzir árvores a partir de um conjunto de treinamento que incorpore valores faltantes, usando os critérios da razão de ganho corrigido (QUINLAN, 1993).

O algoritmo C4.5 apresenta melhorias quando comparado ao ID3, dentre as quais podemos destacar:

- A utilização de poda que remove os ramos que não contribuem para a precisão e os substitue por nós *folha*;
- Permitem o desenvolvimento da indução das árvores embora o conjunto de dados contenha atributos faltantes;
- C4.5 manipula atributos contínuos, dividindo o valor do atributo em análise em dois subconjunto (divisão binária). Em seguida, o algoritmo procura o melhor limiar que maximize o critério da razão de ganho. Todos os valores acima do limiar formarão o primeiro subconjunto e todos os outros valores o segundo.

### 3.3.2.1 Indução de árvores no Algoritmo C4.5

O processo de indução de árvores de decisão tem a função de particionar recursivamente um conjunto de treinamento até que cada subconjunto obtido deste particionamento contenha casos de uma única classe. Uma árvore de decisão toma como entrada um objeto ou situação descrita por um conjunto de atributos e retorna uma decisão com valor de saída previsto, de acordo com a entrada. Os atributos de entrada podem ser discretos ou contínuos (RUSSEL, 2004).

Os algoritmos aplicam o critério de divisão e conquista na construção da árvore, e o critério guloso para escolha de melhores partições e melhores atributos. A árvore de classificação é o resultado de se fazer uma sequência ordenada de perguntas, e as perguntas feitas a cada passo na sequência dependem das respostas às perguntas anteriores. A sequência termina em uma previsão da classe. O ponto de partida de uma árvore de classificação é chamado de nó raiz e consiste em todo o conjunto de aprendizado, e está no topo da árvore. Um nó é um subconjunto do conjunto de atributos, e pode ser terminal ou não-terminal. Um nó não-terminal é um nó que se divide em nós filhos.

Tal divisão é determinada por uma condição sobre o valor de um único atributo, e que vai dividir os exemplos de acordo com a condição, em outros nós. Um nó que não se divide

é chamado de nó terminal, e a ele é atribuída uma classe. Cada exemplo no conjunto cai em um dos nós terminais (IZENMAN, 2008).

O processo de crescimento de uma árvore de classificação, necessita (i) escolher as condições para dividir cada nó, (ii) o critério que deve ser utilizado para dividir um nó pai em seus nós filhos, (iii) a metodologia a ser utilizada para se decidir quando um nó se tornará um nó terminal e (iv) como deve ser atribuída uma classe a esse nó terminal.

Na Figura 23 é ilustrado um exemplo de pseudo-código para a indução de árvores no algoritmo C4.5

Figura 23 – Exemplo de pseudo-código para o C4.5

```

InducaoC4.5(exemplos, subAtributos)
  Se CriterioParada (exemplos)
    EscolheClasse (exemplos).
  Senão
    Melhor = EscolheAtributo (subAtributos, exemplo)
    Arvore = Nova arvore com nó raiz = melhor
    Partição = EscolheParticao (melhor)

    Enquanto partição
      Exp = elementos de exemplos com melhor = p
      subAvr = InducaoC4.5(exp, suba-melhor)
      AdicionarRamoArvore(p, subAvr)

    PODAARVORE(arvore)

  Fim do Enquanto
Fim do Se

```

Fonte – PRÓPRIO AUTOR

Para a etapa *EscolheAtributo*, o critério usado pelo algoritmo para medir a impureza de um nó, pode ser a razão do ganho ou o índice de Gini. Quando este índice é igual a zero, o nó é puro. Por outro lado, quando ele se aproxima de um, o nó é impuro (aumenta o número de classes uniformemente distribuídas neste nó). Quando, o critério de Gini é utilizado tende-se a isolar num ramo os registros que representam a classe mais freqüente.

Na etapa *EscolheParticao*, o algoritmo sempre faz uma divisão binária. Suponha que um determinado atributo é composto por M distintas categorias, o conjunto S de possíveis desdobramentos desse nó é dado por  $(2^{M-1} - 1)$ . A melhor divisão S do nó é dada pela redução da impureza em dividir o nó em filhos direito e esquerdo. A melhor divisão para um atributo, é a

que tem a menor impureza de todas as divisões possíveis ( $2^{M-1} - 1$ ) para o atributo.

Para a etapa *CriterioParada* o algoritmo cresce a árvore até a saturação. Depois de dividir o nó raiz, repetimos o mesmo processo para os nós filhos. Este processo de divisão sequencial de construir uma árvore de camada por camada, é chamado de particionamento recursivo. A árvore é dividida até que nenhum dos nós possa ser dividido mais, até que os nós sejam os mais puros possíveis, ou podemos definir um número mínimo de exemplos para que o nó pare, mas é preferível deixar crescer até a saturação. Pode acontecer dos dados estarem super ajustados ao conjunto de treino, e conter nós com baixa significância estatística, por isso depois de gerada a árvore é necessário podá-la.

Para o método *EscolheClasse*, o algoritmo usa a chamada regra da pluralidade, segundo o qual a classe associada ao nó terminal é a classe que pertence a maioria dos exemplos atribuídos aquele nó. Para o método de *PodaArvore*, o algoritmo utiliza um dos métodos de pós-poda apresentados anteriormente para redução da dimensionalidade da árvore criada.

### 3.4 Considerações Parciais

A utilização do algoritmo C4.5 será a ferramenta chave para o método de diagnóstico utilizado neste trabalho. Inicialmente, o capítulo aborda os conceitos acerca da Descoberta de Conhecimento em Base de Dados, explicando cada uma das 9 etapas do processo.

Posteriormente, foi apresentada a etapa de *Data Mining*, como sendo a de maior importância no processo, por se tratar da etapa que trabalha efetivamente com o tratamento dos dados (muitas vezes inconsistentes) e descobertas das informações ou padrões que se quer alcançar na etapa inicial do processo.

Na seção 3.2 foram apresentados os princípios das Árvores de Decisão, como sendo uma das principais metodologias de Data Mining utilizadas no processo de KDD. Na mesma seção foram abordadas as teorias e procedimentos utilizados para a indução de árvores de decisão bem como a realização da Poda.

Na seção 3.3, foram apresentados os algoritmos ID3 e C4.5, utilizados para indução de árvores de decisão. Destacou-se a interoperabilidade dos citados algoritmos e suas características construtivas. Uma abordagem especial foi dada ao C4.5, visto que este será o algoritmo a ser utilizado no presente trabalho.

No próximo capítulo, será apresentado o desenvolvimento de uma árvore de decisão, utilizando o algoritmo C4.5 para classificação de amostra de gases dissolvidos a partir de dados

cromatográficos extraídos de amostras do óleo de transformadores em operação.

Para a indução da árvore, utilizar-se-á o software *SIPINA*, de uso acadêmico e comercial, específico para mineração de Dados, que implementa vários paradigmas de aprendizagem supervisionado, sendo distribuído gratuitamente pela internet desde 1995.

## 4 UTILIZAÇÃO DE ÁRVORES DE DECISÃO PARA CLASSIFICAÇÃO DGA

Neste capítulo será apresentada a metodologia para a indução da árvore de decisão. A seção 4.1 é destinada ao tratamento da base de dados utilizada, evidenciando o processo de aumento da dimensionalidade do espaço de entrada, técnica desenvolvida em (AMORA et al.2013), sendo utilizada para o desenvolvimento do presente trabalho.

Na seção 4.2 é explicado o processo de indução das árvores de decisão para classificação DGA, fazendo uma análise quanti-qualitativa dos modelos encontrados e das melhorias oferecidas com a evolução do método. Na seção 4.3, já finalizado o processo de indução e escolhida a árvore a ser utilizada serão discutidos os atributos pertencentes ao novo espaço de entradas.

Por fim, na seção 4.4 serão feitas as considerações parciais acerca do capítulo 4.

### 4.1 Seleção e tratamento de dados

Conforme descrito na *seção 2.5.3*, utilizou-se para o desenvolvimento do presente trabalho um banco de dados conhecido no meio acadêmico e que auxiliam pesquisas voltadas para o diagnóstico de falhas incipientes em transformadores de força.

Foram utilizadas 162 amostras contendo resultados de ensaios cromatográficos em óleo de transformadores de força em operação. Cada amostra contém primariamente 8 atributos, sendo 7 características relacionadas aos gases dissolvidos e 1 correspondente a classificação da falta. Os atributos são (em ppm):

- Hidrogênio( $H_2$ );
- Metano( $CH_4$ );
- Acetileno( $C_2H_2$ );
- Etileno( $C_2H_4$ );
- Etano( $C_2H_6$ );
- Monóxido de carbono( $CO$ );
- Dióxido de carbono( $CO_2$ );

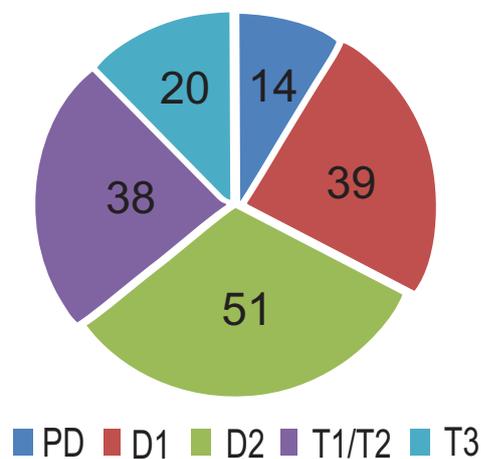
O oitavo atributo de cada amostra é o tipo de falta incipiente associada à amostra em análise, que pode ser descrita como:

- PD (Descargas parciais);
- D1 (Falha elétrica de baixa energia);

- D2 (Falha elétrica de alta energia);
- T1/T2 (Falha térmica envolvendo temperaturas que variam até 700°C);
- T3 (Falha térmica envolvendo temperaturas acima de 700°C);

Das 162 amostras, 117 foram extraídas de (IEC 60599, 2008), 39 das Tabelas de I a III de (DUVAL, 2002) e 6 amostras de transformadores em operação apresentadas em (NARESH, 2008). A distribuição das classes levando em consideração o tipo de falta incipiente é ilustrada na Figura 24.

Figura 24 – Espaço amostral



Fonte – PRÓPRIO AUTOR

Conforme apresentado em Amora (2013), todas as amostras são submetidas ao processo de aumento da dimensionalidade do espaço de entrada antes de iniciarem a etapa de pré-processamento e limpeza dos dados.

#### **4.1.1 Aumento da dimensionalidade do espaço de entrada**

(Cover et al. 1965) apresenta um conjunto de teoremas e definições fundamentais relacionados com a solução de problemas complexos de classificação de padrões, demonstrando que tais problemas quando dispostos em um espaço de alta dimensão apresentam uma maior probabilidade de ser linearmente separáveis do que em um espaço de baixa dimensionalidade (AMORA, 2013).

A metodologia desenvolvida em AMORA (2013) propõe a solução de problemas de

classificação inicialmente realizando um aumento das variáveis do problema, utilizando novos atributos originalmente não utilizados e relações entre os atributos do espaço de entrada original. Com isso, é obtida uma nova representação do problema em uma dimensão maior e, conforme indicado em Cover (1965), o problema passa a ter uma probabilidade maior de levar os padrões do novo espaço de atributos a serem linearmente separáveis, e a uma divisão mais favorável desse novo espaço de entrada por hiperplanos.

Os novos atributos propostos devem ser linearmente independentes dos atributos originais. Nesta pesquisa, os atributos incorporados no modelo do problema, além de informações novas originalmente não utilizadas, são relações não lineares entre os atributos originais. Dessa forma, se for considerado como atributos originais  $x_1$  e  $x_2$ , os novos atributos propostos têm os formatos de  $x_3$  e  $x_4$ , sendo apresentados nas equações (4.1) e (4.2) por exemplo.

$$x_3 = x_1 \cdot x_2 + b \quad (4.1)$$

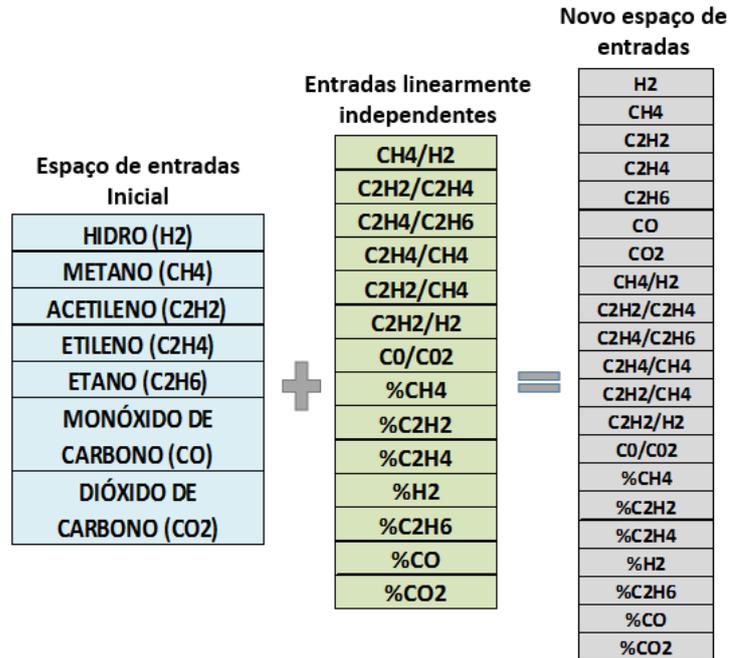
$$x_4 = \frac{x_1}{x_2 + b} \quad (4.2)$$

Além de novos atributos com base nas relações acima, são também agregados ao espaço de entrada, atributos que já são utilizados em outros métodos que utilizam razões como Rogers e Dornembürg.

Dessa forma, o novo espaço de entradas aumentado, tomando por base relações linearmente independentes entre os atributos existentes que são incorporados ao problema de classificação, estão descritos na Figura 25.

Este conterá 21 atributos por amostra sendo que destas, 7 são concentrações em *ppm* dos gases extraídos diretamente da amostra cromatográfica sem nenhum tipo de tratamento, que são: 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( $CO_2$ ). Para composição do espaço, utilizou-se também as relações:  $CH_4/H_2$ ,  $C_2H_2/C_2H_4$ ,  $C_2H_4/C_2H_6$ ,  $C_2H_2/CH_4$ , que são relações advindas dos métodos de análise DGA que utilizam razões como Rogers e Dornembürg,  $C_2H_4/CH_4$ ,  $C_2H_2/H_2$  e a razão  $CO/CO_2$  que é a relação inversa da utilizada em (IEEE, 2008) para medir o grau de degradação do papel devido temperaturas excessivas.

Figura 25 – Novo Espaço de entrada



Fonte – PRÓPRIO AUTOR

Seguindo as premissas adotadas por DUVAL ao utilizar relações percentuais de gases, %C<sub>2</sub>H<sub>2</sub>, %C<sub>2</sub>H<sub>4</sub>, %CH<sub>4</sub>, como atributos de entrada para o seu método, optou-se por inserir como atributos as concentrações relativas abaixo:

$$\%H_2 = \frac{100 \cdot [H_2]}{[H_2] + [C_2H_6] + [CO] + [CO_2]} \quad (4.3)$$

$$\%C_2H_6 = \frac{100 \cdot [C_2H_6]}{[C_2H_6] + [CH_4] + [C_2H_2] + [C_2H_4]} \quad (4.4)$$

$$\%CO = \frac{100 \cdot [CO]}{[C_2H_6] + [CH_4] + [C_2H_2] + [C_2H_4] + [CO] + [CO_2]} \quad (4.5)$$

$$\%CO_2 = \frac{100 \cdot [CO_2]}{[C_2H_6] + [CH_4] + [C_2H_2] + [C_2H_4] + [CO] + [CO_2]} \quad (4.6)$$

Na próxima seção, são apresentadas as considerações adotadas na etapa de pré-processamento de dados, bem como o desenvolvimento do algoritmo para formação das árvores de decisão objeto deste trabalho.

#### 4.1.2 Pré-processamento e limpeza dos dados

Para a etapa de pré-processamento e higienização dos dados, buscando assim eliminar inconsistências existentes na base, bem como recuperação de dados incompletos, foram realizadas as seguintes considerações:

- Se a concentração de determinado gás não estiver disponível, então deve-se considerar esta igual a zero;
- Uma razão definida por  $0/0$  é considerada nula (LI e WU,2011);
- Uma razão definida como  $v/0$  será ajustada para 20, considerando que  $v$  não é nulo (LI e WU,2001);
- Uma concentração indicada como sendo  $<1$  deverá ter o valor de 0.5 (LI e WU,2011).

Uma vez aplicada as considerações acima, o banco de dados está apto a ser apresentado para o algoritmo C4.5 afim de iniciar o processo de indução das árvores de decisão.

Para o início do processo de indução da árvore, o algoritmo deve escolher o atributo de maior relevância, ou seja, aquele que é capaz de dividir melhor o banco de dados, consequentemente, tenha a maior *razão do ganho* e a menor *entropia* inserida no conjunto do citado atributo.

Para os 21 atributos (vide Figura 25) apresentados ao software *SIPINA*, o algoritmo selecionou o atributo  $PC2H2(\%C_2H_2)$  como sendo o nó raiz da árvore.

A escolha da concentração relativa de  $C2H2(\%C_2H_2)$ , corrobora com a escolha feita por DUVAL na concepção da sua metodologia, como sendo esse atributo um dos lados do triângulo. Este por sua vez, é fundamental pois traz informações não apenas da concentração de acetileno, mais também do metano( $CH_4$ ) e do etileno( $C_2H_4$ ), que são os 3 gases combustíveis mais nocivos ao funcionamento do transformador de força, e que sua detecção em uma análise cromatográfica, para determinadas quantidades, pode denotar falhas incipientes de alta relevância.

Na Tabela 7 é apresentada a análise realizada pelo algoritmo C4.5 para a escolha do atributo que ocupará o nó da árvore de decisão. Pela análise dos valores, o melhor atributo a ser utilizado para a divisão inicial da árvore é o  $PC2H2(\%C_2H_2)$ , com o valor de 0.916658 para a razão do ganho, seguido pelo atributo  $PCH4(\%CH_4)$  com o valor de 0.750513.

Ainda com base na Tabela 7, observa-se que os parâmetros advindos do triângulo de DUVAL, estão entre os 4 melhores atributos para início da árvore. Tal fato a priori, não gera surpresa visto que a metodologia de DUVAL é bastante eficaz tanto no meio acadêmico como na

Tabela 7 – Escolha do Atributo para o Nó Raiz da Árvore de Decisão

Atributo	Razão do Ganho	Correlação
%C2H2	0.91665804	0.47
%CH4	0.75051333	0.37
C2H2/C2H4	0.59473788	0.32
%C2H4	0.59251079	0.31
H2	0.55033254	0.20
C2H2/CH4	0.53078195	0.29
C2H2	0.49587454	0.28
C2H2/H2	0.49256932	0.22
C2H4/C2H6	0.42430540	0.22
CH4/H2	0.36901970	0.21
C2H4	0.36851092	0.20
%H2	0.29445029	0.18
C2H4/CH4	0.28105414	0.16
%C2H6	0.26108098	0.15
CO2	0.20785867	0.06
%CO2	0.17232424	0.05
CO	0.17216752	0.05
C2H6	0.15791294	0.06
CO/CO2	0.14976980	0.03
%CO	0.13274231	0.08
CH4	0.12725088	0.08

utilização em departamentos de manutenção em concessionárias de energia elétrica.

Generalizando, pode-se inferir que a maior quantidade de informações acerca de falhas incipientes em transformadores é norteadada pela análise de 4 atributos, sendo eles o metano, etileno, acetileno e hidrogênio.

Os demais atributos, como o etano, o monóxido e o dióxido de carbono, tem sua importância para caracterizar o tipo de falha, refinando o processo de busca na classificação DGA, sendo que os dois últimos, são relacionados a degradação da celulose na ocorrência de faltas, que envolvam o papel isolante do enrolamento do transformador.

Para o atributo %C2H2, escolhido como nó-raiz pelo algoritmo, a divisão das amostras será realizada conforme Tabela 8.

Para a escolha do atributo que será o segundo estágio da árvore, o algoritmo calcula novamente a *razão do ganho*, levando em consideração todos os parâmetros disponíveis, inclusive o atributo PC2H2 utilizado como nó da árvore. Cabe ressaltar que, por se tratar de um algoritmo que tem a busca "gulosa", este irá procurar o atributo que forneça o maior ganho local,

Tabela 8 – Divisão das Amostras pelo Nó-raiz

	<12,89	>=12,89
PD	14	0
D1	1	38
D2	0	51
T1/T2	37	1
T3	20	0

independente da continuidade da árvore e da utilização recorrente ou não do atributo escolhido.

A Tabela 9, apresenta a análise realizada pelo algoritmo C4.5 para a escolha do 2º nó da árvore.

Tabela 9 – Escolha do Atributo para o 2º Nó Raiz da Árvore de Decisão

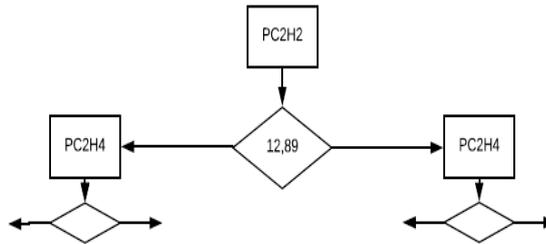
Atributo	Razão do Ganho	Correlação
%C <sub>2</sub> H <sub>4</sub>	0.73541711	0.39
C <sub>2</sub> H <sub>4</sub> /CH <sub>4</sub>	0.61837275	0.34
%CH <sub>4</sub>	0.58023699	0.30
C <sub>2</sub> H <sub>2</sub> /C <sub>2</sub> H <sub>4</sub>	0.53706057	0.18
%H <sub>2</sub>	0.47235422	0.28
C <sub>2</sub> H <sub>4</sub>	0.42695639	0.22
%CO	0.33643583	0.13
C <sub>2</sub> H <sub>4</sub> /C <sub>2</sub> H <sub>6</sub>	0.30676980	0.18
H <sub>2</sub>	0.28599955	0.17
CH <sub>4</sub> /H <sub>2</sub>	0.28031194	0.16
C <sub>2</sub> H <sub>2</sub>	0.26647514	0.16
CO <sub>2</sub>	0.23936354	0.14
%C <sub>2</sub> H <sub>2</sub>	0.21925323	0.14
CH <sub>4</sub>	0.19768305	0.12
C <sub>2</sub> H <sub>2</sub> /H <sub>2</sub>	0.18231603	0.11
C <sub>2</sub> H <sub>2</sub> /CH <sub>4</sub>	0.18231603	0.11
%C <sub>2</sub> H <sub>6</sub>	0.13165170	0.08
%CO <sub>2</sub>	0.1032528	0.05
C <sub>2</sub> H <sub>6</sub>	0.09702177	0.05
CO	0.00000000	0.00
CO/CO <sub>2</sub>	0.00000000	0.00

Para a composição do 2º nó, o algoritmo C4.5 utilizou o atributo PC<sub>2</sub>H<sub>4</sub>(%C<sub>2</sub>H<sub>4</sub>). Nota-se que uma vez escolhido o primeiro atributo (nó raiz), este momentaneamente passa para um patamar inferior de prioridade, permitindo que demais atributos continuem a divisão da árvore. A ideia é que divisão após divisão, nó após nó, seja possível visualizar os agrupamentos

entre as amostras iguais, uma vez que para o algoritmo C4.5 a árvore cresce sem nenhuma medida de controle de tamanho.

Após a composição do 2º estágio, tem-se a árvore de decisão induzida conforme apresentado na Figura 26.

Figura 26 – 1º e 2º Nós da Árvore de Decisão



Fonte – PRÓPRIO AUTOR

Para a escolha dos demais nós, a metodologia a ser utilizada será a mesma, até que não exista mais possibilidades (com base nos atributos disponíveis) de se realizar as divisões necessárias para separação de amostras.

A seção seguinte apresentará as árvores induzidas pelo algoritmo C4.5 em sua forma completa, bem como uma análise quanti-qualitativa acerca de cada uma destas.

## 4.2 Árvores de decisão para classificação DGA

Para a escolha da árvore com melhor acurácia na classificação DGA, optou-se pela indução de 5 árvores utilizando o Software *SIPINA*. Para todos os processos de treinamento/teste, utilizou-se a validação cruzada, dividindo-se em 10 grupos as 162 amostras, sendo que para cada etapa de treinamento eram destinados 9 grupos e o remanescente era utilizado como teste. Tal procedimento é utilizado quando não se dispõe de uma base de dados significativamente grande.

### 4.2.1 Árvore 01

Para a primeira árvore, apresentou-se as 162 amostras com os 21 atributos ao algoritmo C4.5. Com a construção da referida árvore, observou-se que esta seria a de menor dimensão (29 nós) porém a de menor acerto (98,15%). Observou-se ainda que a construção da árvore pelo algoritmo utilizou apenas 10 dos 21 atributos apresentados, classificando os demais como não fundamentais para o processo de classificação DGA.

A Figura 27 ilustra a Árvore 1, na qual os nós de decisão são representados pelos retângulos com o seu respectivo atributo. As quantidades representadas nos losangos correspondem aos valores limítrofes para a divisão da amostra com base no atributo local e as circunferências correspondem aos nós-folhas com as classes correspondentes a classificação das amostras.

Ainda pela análise da referida árvore, o símbolo "<" significa MENOR QUE o valor do atributo considerado localmente, e o símbolo ">=" significa MAIOR OU IGUAL, em relação ao valor atributo em questão. As setas indicam o fluxo da informação após a averiguação local.

A Tabela 10 apresenta a matriz de confusão para classificação DGA das 162 amostras utilizando a Árvore 1.

Tabela 10 – Matriz de Confusão para classificação DGA utilizando a árvore 1

	<b>PD</b>	<b>D1</b>	<b>D2</b>	<b>T1/T2</b>	<b>T3</b>
<b>PD</b>	14	0	0	0	0
<b>D1</b>	0	38	0	0	1
<b>D2</b>	0	1	50	0	0
<b>T1/T2</b>	1	0	0	37	0
<b>T3</b>	0	0	0	0	20

Pela análise da matriz de confusão após a classificação feita pela primeira árvore verifica-se que o algoritmo classificou de forma incorreta 03 amostras:

- Amostra *D1* classificada como *T3*;
- Amostra *D2* classificada como *D1*;
- Amostra *T1/T2* classificada como *PD*;

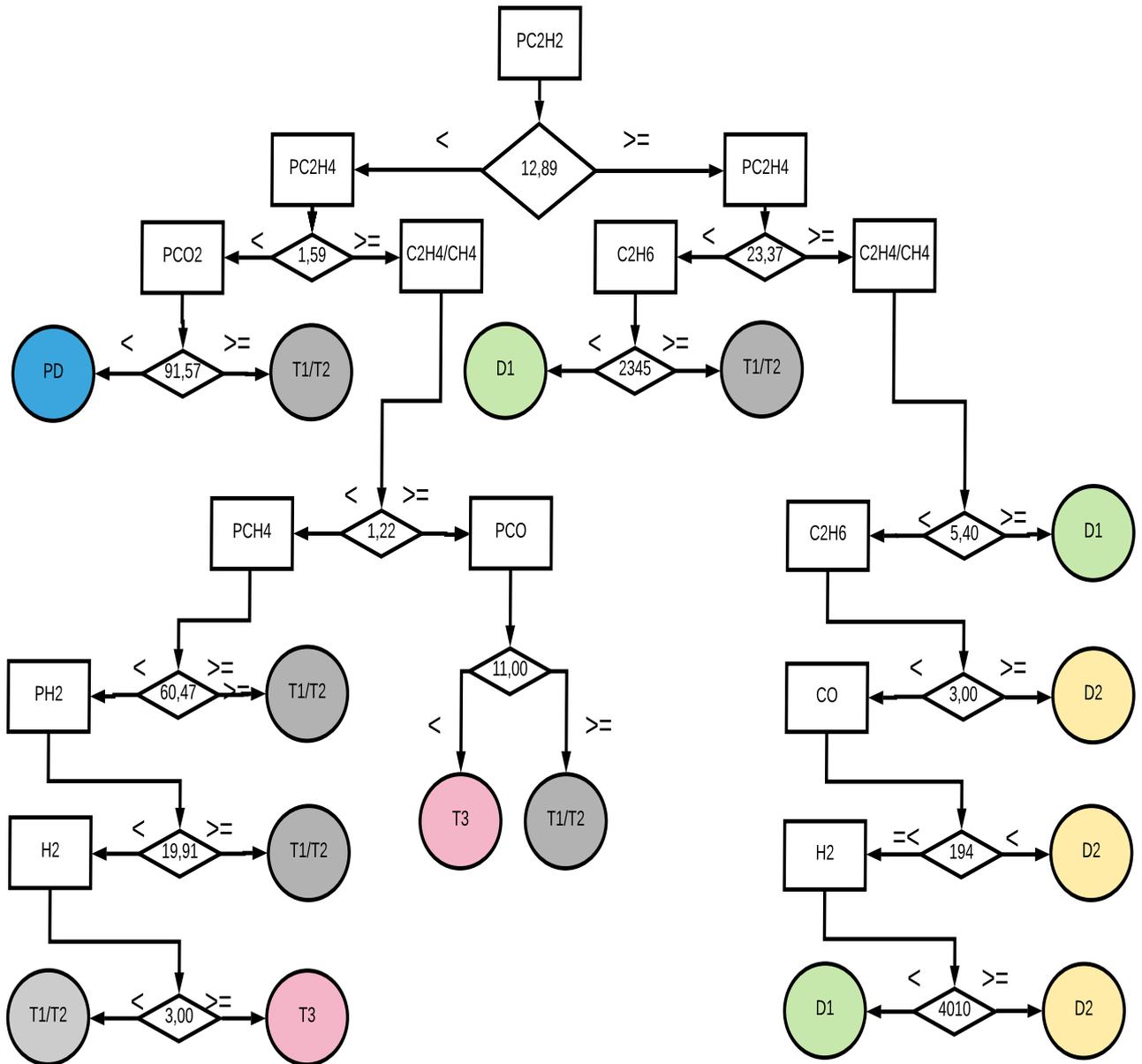
Para a amostra *D1*, classificada como *T3*, o erro ocorreu no nó *PCO*, na decisão para o valor de 11,00 (vide Figura 28).

Para a amostra *D2*, classificada como *D1*, o erro ocorreu no nó *C2H6*, na decisão para o valor de 2345,00 (vide Figura 29).

Para a amostra *T1/T2*, classificada como *PD*, o erro ocorreu no nó *PCO2*, na decisão para o valor de 91,57 (vide Figura 30).

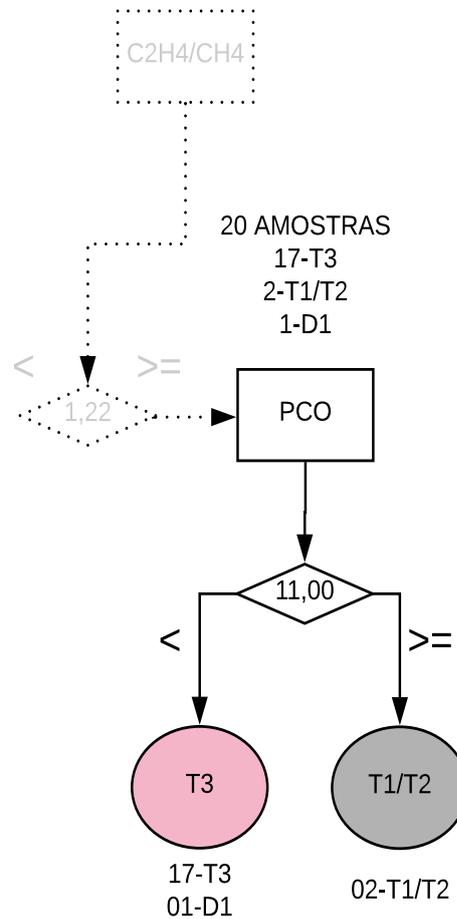
Ao se realizar uma análise detalhada, atributo à atributo das amostras classificadas como erradas pela árvore 01 observa-se que para o caso da amostra *D1*, esta tem todas as características de uma amostra *T3* e a amostra *T1/T2* encaixa-se perfeitamente como uma amostra *PD*.

Figura 27 – Árvore 1



Fonte – PRÓPRIO AUTOR

Figura 28 – Erro 01 árvore 1



Fonte – PRÓPRIO AUTOR

A amostra D2, classificada como D1, pelo método do triângulo de Duval está localizada na região limítrofe entre as duas regiões de classificação.

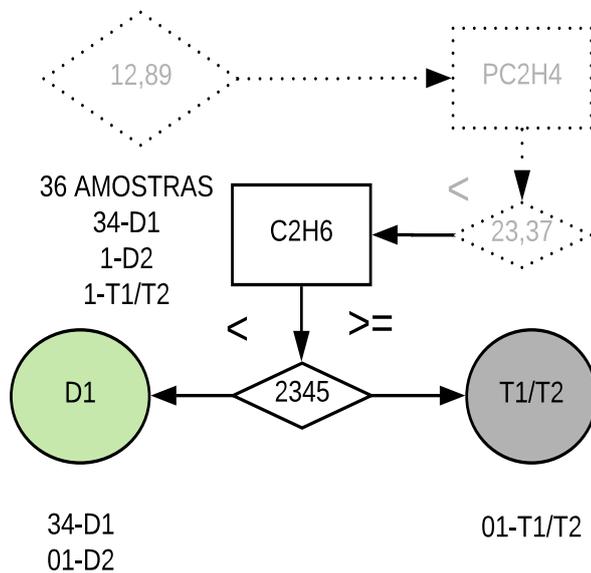
#### 4.2.2 *Árvore 02*

Para a árvore 2, foram retirados os atributos %CO e %H2 (mapeados na árvore 1), devido a presença destes atributos juntamente com outros de mesmo significado físico na árvore (CO e H2) obtendo-se com isso uma árvore de dimensão maior (33 nós), porém com um acerto de 98,77%. A Figura 31 traz a representação da Árvore 2.

A Tabela 11 apresenta a matriz de confusão para classificação DGA das 162 amostras utilizando a Árvore 2.

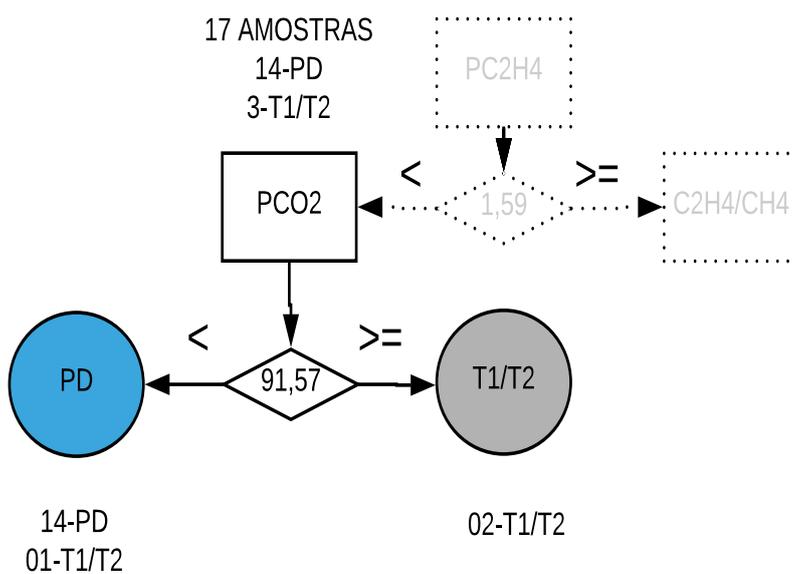
Pela análise da matriz de confusão após a classificação feita pela segunda árvore

Figura 29 – Erro 02 Árvore 1



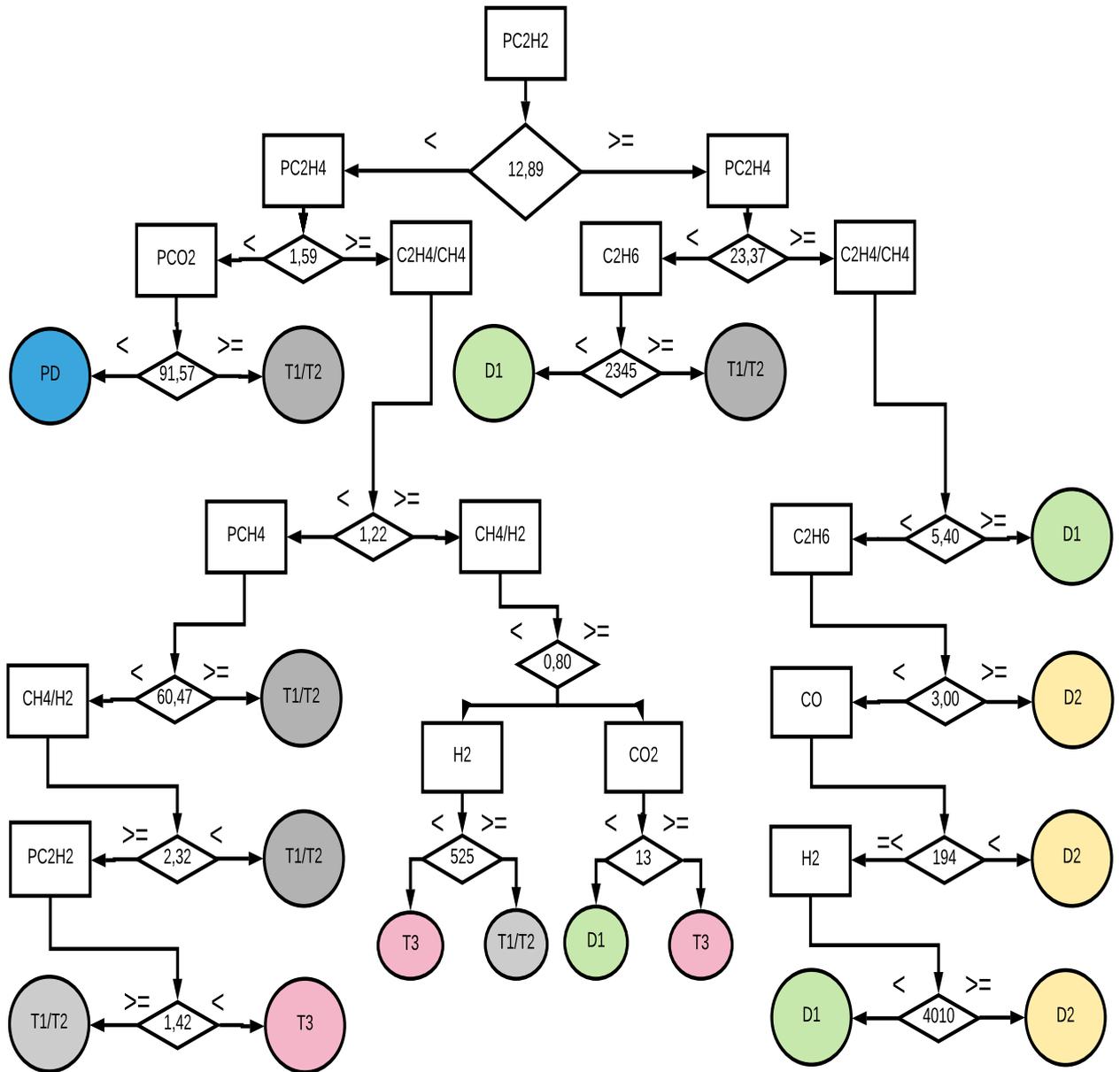
Fonte – PRÓPRIO AUTOR

Figura 30 – Erro 03 Árvore 1



Fonte – PRÓPRIO AUTOR

Figura 31 – Árvore 2



Fonte – PRÓPRIO AUTOR

Tabela 11 – Matriz de Confusão para classificação DGA utilizando a árvore 2

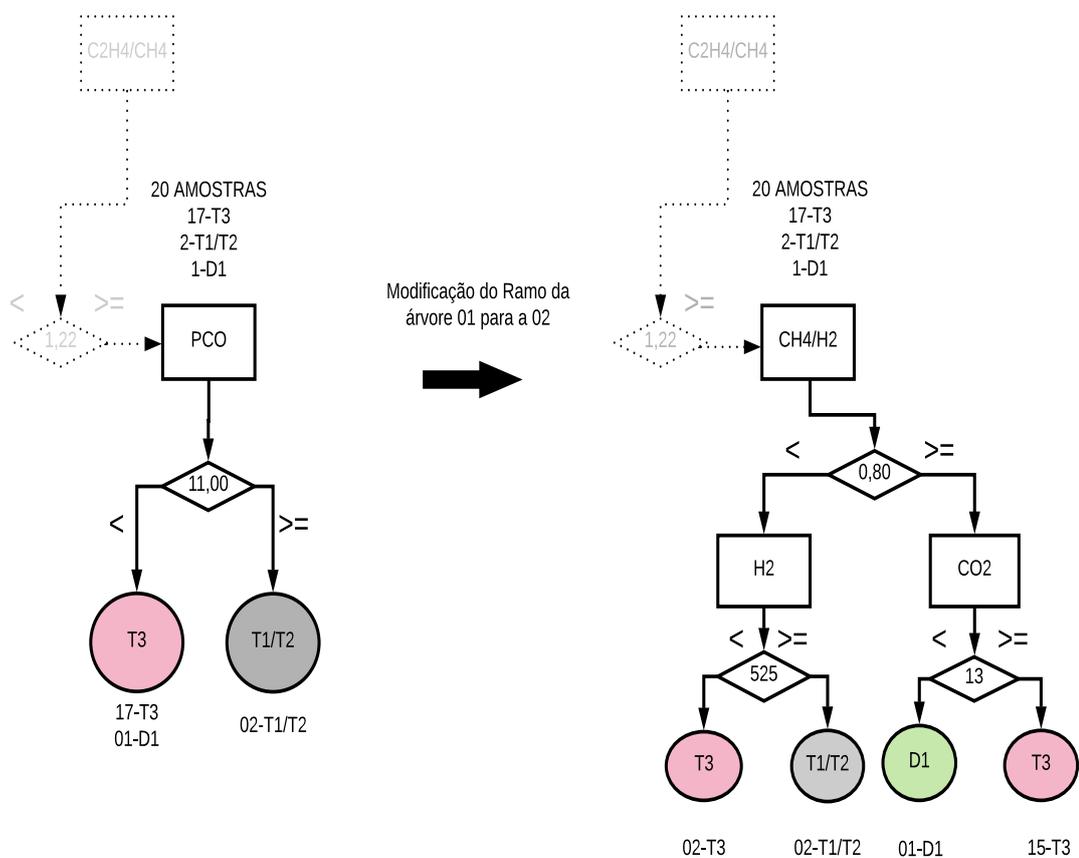
	PD	D1	D2	T1/T2	T3
PD	14	0	0	0	0
D1	0	39	0	0	0
D2	0	1	50	0	0
T1/T2	1	0	0	37	0
T3	0	0	0	0	20

verifica-se que o algoritmo classificou de forma incorreta 02 amostras:

- Amostra *D2* classificada como *D1*;
- Amostra *T1/T2* classificada como *PD*;

Na indução da segunda árvore, o erro de classificação ocorrido na amostra *D1* classificada como *T3* pela primeira árvore foi suprimido conforme a Figura 32 e o ramo responsável pela citada classificação foi substituído.

Figura 32 – Modificação ocorrida da árvore 01 para a 02



Fonte – PRÓPRIO AUTOR

Para o aumento no percentual de acerto apresentado pela árvore 02 em relação a árvore 01 (de 98,15% para 98,77%), foi necessário a inserção de 4 nós a mais na árvore, passando esta de 29 para 33 nós. Computacionalmente, esse aumento não gera custo, visto que o tempo de processamento das duas árvores, embora a segunda tenha uma maior dimensionalidade não ultrapassa poucos milésimos de segundos.

### 4.2.3 Árvores 03, 04 e 05

Para as árvores 3, 4 e 5 independentes dos atributos apresentados ao algoritmo, porém permanecendo %C<sub>2</sub>H<sub>2</sub>, %C<sub>2</sub>H<sub>4</sub>, %CH<sub>4</sub>, H<sub>2</sub>, C<sub>2</sub>H<sub>6</sub>, CO, CO<sub>2</sub>, CH<sub>4</sub>, C<sub>2</sub>H<sub>4</sub>/CH<sub>4</sub>, CH<sub>4</sub>/H<sub>2</sub> a dimensão da árvore permanece em 35 nós, 18 folha (regras) e com uma capacidade de acerto de 99,38%, classificando corretamente 161 das 162 amostras apresentadas. A Figura 33 traz a representação da árvore final.

A Tabela 12 apresenta a matriz de confusão para classificação DGA das 162 amostras utilizando a Árvore Final.

Tabela 12 – Matriz de Confusão para classificação DGA utilizando a árvore final

	PD	D1	D2	T1/T2	T3
PD	14	0	0	0	0
D1	0	39	0	0	0
D2	0	1	50	0	0
T1/T2	0	0	0	38	0
T3	0	0	0	0	20

Pela análise da matriz de confusão após a classificação feita pela terceira árvore, verifica-se que o erro encontrado foi a classificação de uma amostra D2 como D1.

Na indução da citada árvore, o erro de classificação ocorrido na amostra T1/T2 classificada como PD pelas duas primeiras árvores foi suprimido conforme a Figura 34 e o ramo responsável pela citada classificação foi substituído.

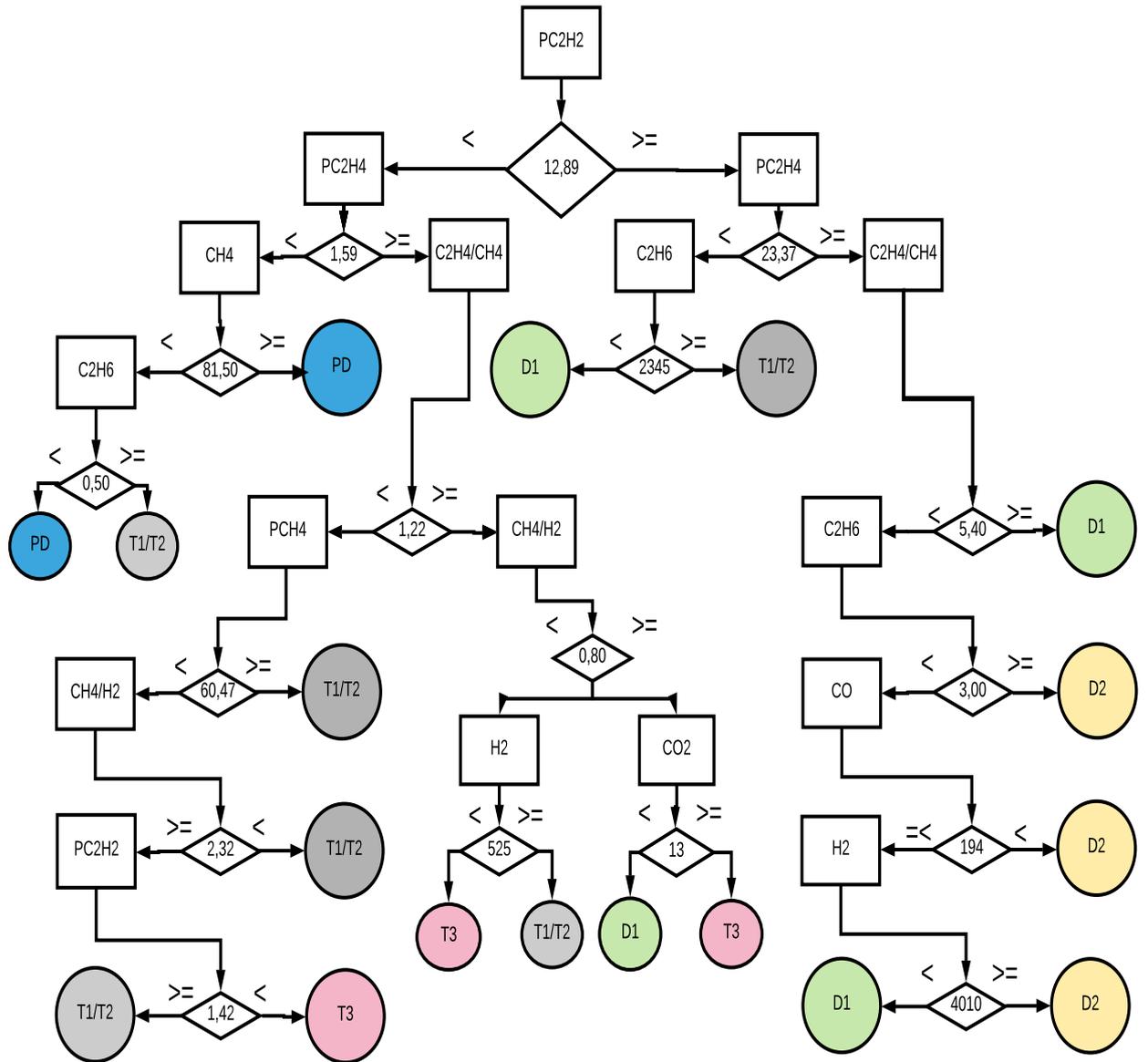
As árvores 4 e 5 acompanham a mesma estrutura da árvore 03, ou seja, o algoritmo C4.5 alcança com a indução da referida árvore o seu ponto ótimo, sendo indispensáveis quaisquer manipulações posteriores afim de melhorar o índice de 99,38% de acerto.

Similar a uma rede neural do tipo MLP, em que se acrescentam neurônios na camada intermediária afim de melhorar o acerto na classificação e estabilização do erro médio, a árvore de decisão atinge com essa configuração seu limiar de acerto e estabilidade no erro, não bastando para tanto a inserção de mais nós ao algoritmo para fins de melhoria de performance.

### 4.3 Composição do novo espaço de entradas

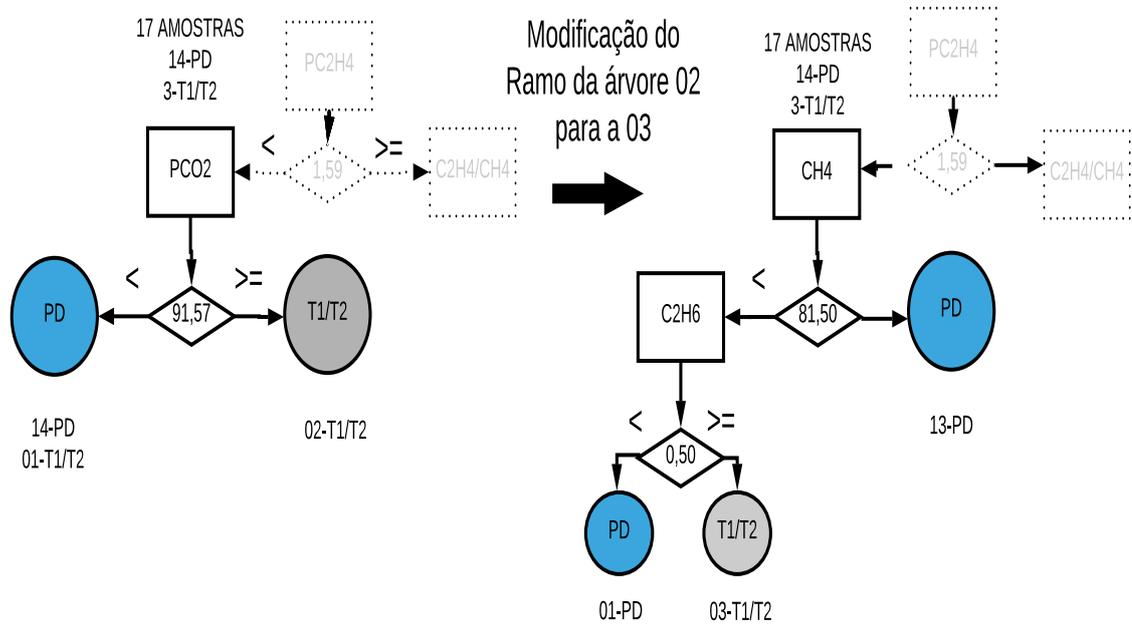
Para a composição da árvore final, tem-se os seguintes atributos para o novo espaço de entradas, conforme Tabela 13:

Figura 33 – Árvore Final



Fonte – PRÓPRIO AUTOR

Figura 34 – Modificação ocorrida da árvore 2 para a 3



Fonte – PRÓPRIO AUTOR

Tabela 13 – Novo espaço de entradas para classificação DGA

Atributo	Origem
%C2H2 %C2H4 %CH4	Triângulo de Duval
H <sub>2</sub> CH <sub>4</sub> CO CO <sub>2</sub> C <sub>2</sub> H <sub>6</sub>	Atributos primário
C <sub>2</sub> H <sub>4</sub> /CH <sub>4</sub>	Regras extraídas(AMORA,2013)
CH <sub>4</sub> /H <sub>2</sub>	Rogers e Dornenburg

Por *Atributo primário*, entende-se aquele que é adquirido diretamente da amostra cromatográfica, sem nenhum tipo de manipulação posterior seja para compor uma relação ou proporção.

Denota-se pela análise da Tabela 13, a contribuição de cada um dos métodos propostos até então para a consolidação de um trabalho com resultados satisfatórios, desde os estudos de *Rogers* até as regras extraídas em *Amora* (2013).

#### 4.4 Considerações Parciais

O processo de classificação de qualquer banco de dados inicia-se com a etapa de seleção e tratamento das amostras a serem utilizadas. Neste capítulo, inicialmente mostrou-se a forma com que foi composto o novo espaço de entradas, a partir da escolha e tratamento dos atributos a serem utilizados.

Demostrou-se também, a forma que o algoritmo C4.5, através do software *SIPINA*, trabalha na escolha do atributo inicial e conseqüentemente dos demais atributos, buscando sempre o que representa uma melhor repartição das amostras, tendo assim, o atributo a ser escolhido, a maior *razão do ganho* e conseqüentemente a menor *entropia* local.

Na seção 4.2 foram induzidas 5 árvores de decisão utilizando o software *SIPINA*. Foi demonstrado a evolução do algoritmo, árvore após árvore, até se alcançar o ponto ideal de parada, de onde não seria mais possível agregar resultado com o aumento da dimensionalidade da árvore proposta.

Por fim, na seção 4.3 foram apresentados os atributos que foram utilizados para a composição da árvore de decisão final, formatada pelo algoritmo C4.5 para classificação DGA, bem como a procedência de cada um destes.

No próximo capítulo, serão apresentados os resultados obtidos, bem como as novas regras extraídas para a classificação de falhas incipientes de transformadores utilizando árvores de decisão. Serão também apresentados, os significados práticos de cada um dos dez atributos pertencentes ao novo espaço de entradas na classificação DGA pelo método proposto neste trabalho.

## 5 RESULTADOS

Neste capítulo são apresentadas as considerações acerca da utilização de árvores de decisão para a classificação DGA.

Na seção 5.1, efetua-se um detalhado estudo acerca de cada um dos atributos utilizados pela DT proposta, afim de localizarmos quais atributos são utilizados na detecção de determinado tipo de falta.

Na seção 5.2, apresentam-se as regras extraídas para a classificação DGA, após a indução da árvore de decisão final. Na seção 5.3 será apresentado uma nova proposta para o triângulo de Duval, de maior acurácia, a partir de regras extraídas a partir de uma árvore de decisão.

Durante a seção 5.4 é apresentada a proposta de uma regra para classificação de amostras em térmicas ou elétricas, chamada de *Regra do Acetileno*, e na seção 5.5 serão validados os métodos propostos até então, utilizando casos de análises cromatográficas presente em trabalhos acadêmicos atuais.

Por fim, a seção 5.6 apresenta as considerações parciais acerca do presente capítulo.

### 5.1 Rastreabilidade dos Atributos Seleccionados para Composição da Árvore de Decisão

O resultado final do processo de indução de árvores de decisão é a extração de regras disjuntivas sequenciais, que quando utilizadas, resultarão na classificação da amostra em um dos tipos de falhas indicadas na seção 4.1.

Uma outra forma de se analisar os atributos seleccionados para a construção da árvore, é a análise de qual (ou quais) atributos estão presentes na classificação de certo tipo de falta, ou seja, quais são os atributos que quando de sua ocorrência remetem determinada condição indicativa de uma condição específica.

- Percentual de Acetileno( $\%C_2H_2$ ): É o nó principal da DT. Sua concentração inicial divide as amostras em grupos de falhas elétricas e térmicas. Em geral, para amostras cujo parâmetro ( $\%C_2H_2$ ) é menor do que 12,89 (vide Figura 35), descarta-se a ocorrência de falhas elétricas, ficando o algoritmo responsável pela classificação destas em DP, T1/T2 ou T3. Comparando-se a preposição encontrada com a teoria de formações dos gases(IEEE,2008), observa-se a coerência com o resultado apresentado, uma vez que o acetileno apresenta traços a partir

de 500°C, temperatura característica do início de formação de falhas de natureza elétrica internas ao transformador.

Pela análise da Figura 36, observa-se o início da formação do gás acetileno em poucas quantidades a partir de 500°C, sendo que sua presença acima de 700°C é o maior indicativo de que falhas elétricas ocorrem internamente, fato este mapeado pela árvore de decisão no nó raiz selecionado:

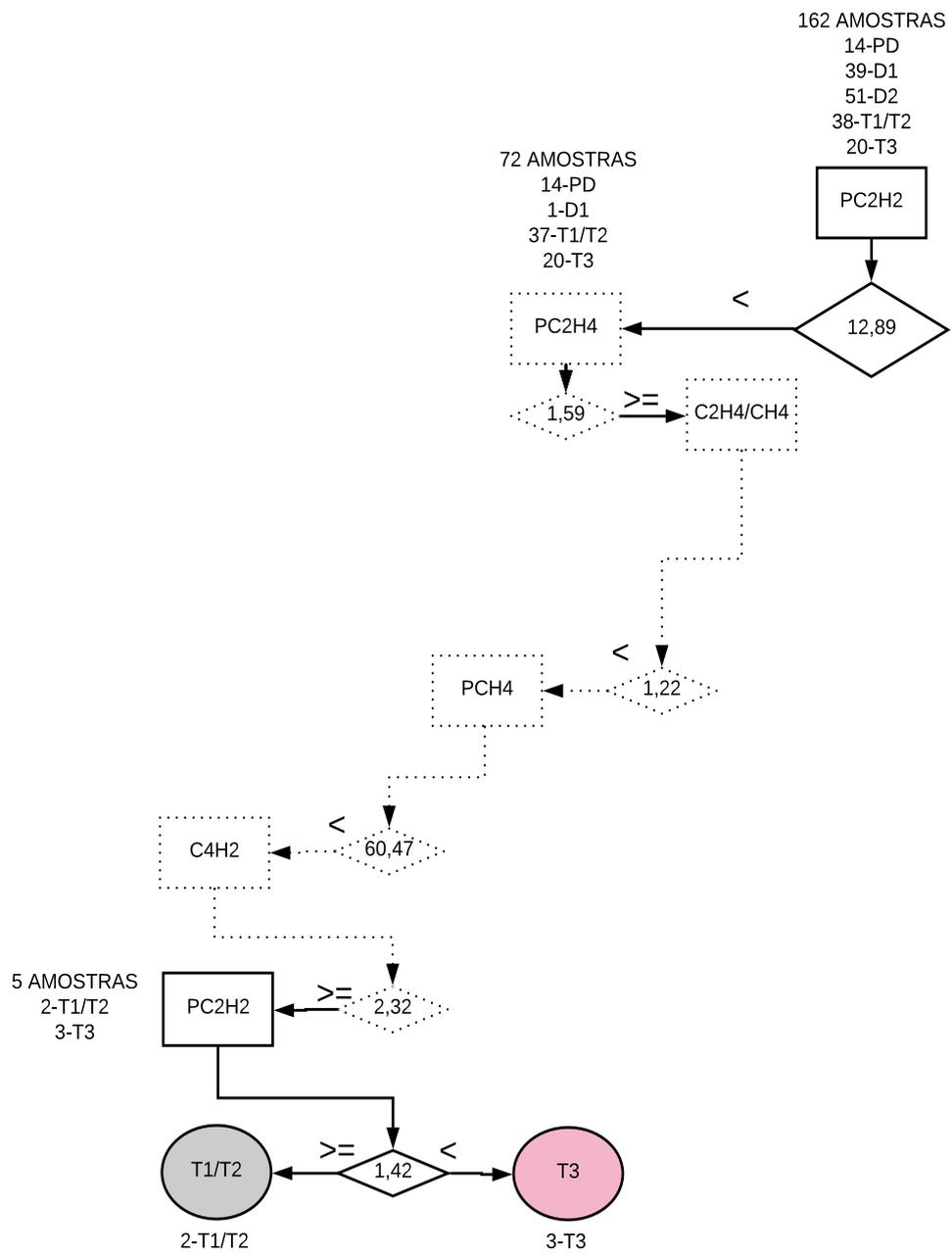
- Percentual de Etileno ( $\%C_2H_4$ ): É o atributo presente no segundo nível da DT. A Figura 37 mostra que para o limiar de decisão em 23,37 atua na divisão das falhas elétricas em D1 e D2. No limiar de 1,59 divide as amostras PD das falhas térmicas T1/T2 e T3. Em um processo de manutenção preditiva com acompanhamento da evolução das quantidades de gases gerados no interior do transformador, pode-se inferir que a evolução do valor de ( $\%C_2H_4$ ), ultrapassando o limite acima, remete a evolução de uma falha elétrica que deve ser imediatamente verificada pela equipe de manutenção, para que assim possam ser evitadas avarias maiores no equipamento.

Realizando a comparação com o gráfico da teoria de formação dos gases na Figura 38, observa-se que a árvore de decisão para os ramos do atributo etileno atua na divisão tanto para baixas temperaturas, separando as falhas térmicas quanto em altas temperaturas atuando na separação das falhas elétricas. Observa-se mais uma vez que, conforme apontado pela árvore proposta, em baixas concentrações esta divide T1/T2 de T3 e em concentrações maiores, atua na divisão de D1 em D2, associado a concentrações maiores de acetileno.

- Percentual de Metano ( $\%CH_4$ ): Atributo presente no ramo da árvore cujo valor de acetileno ( $\%C_2H_2$ ) é menor do que 12,89 (vide Figura 39). Associado ao atributo  $C_2H_4/CH_4$ , atua principalmente na classificação das falhas térmicas T1/T2, sendo sua presença em valores maiores que 60,47 um forte indicativo para a ocorrência de tal anomalia.

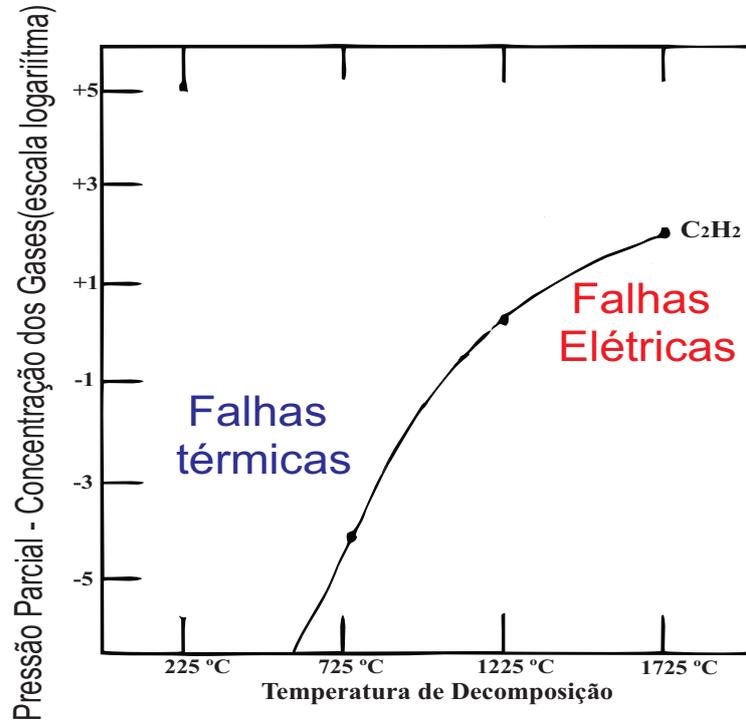
A comprovação física da proposta indicada para o atributo metano pela árvore de decisão é ratificada quando se utiliza novamente a teoria de formação dos gases (IEEE,2008). Pela análise da Figura 40, observa-se que para valores baixos de acetileno (ramo da árvore para valores de acetileno menores que 12,89), existe

Figura 35 – Rastreabilidade da porcentagem de acetileno ( $PC_2H_2$ )



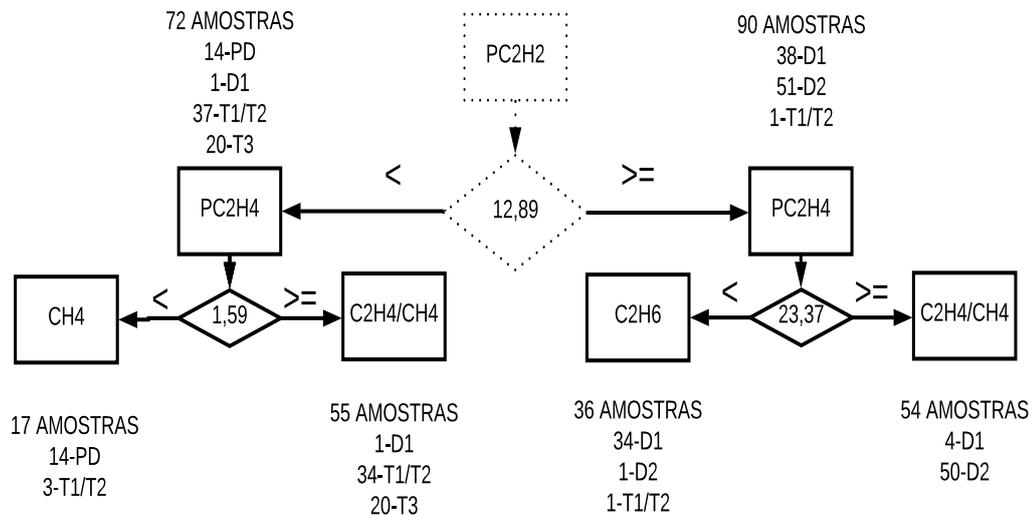
Fonte – PRÓPRIO AUTOR

Figura 36 – Divisão do acetileno segundo a teoria de formação dos gases dissolvidos



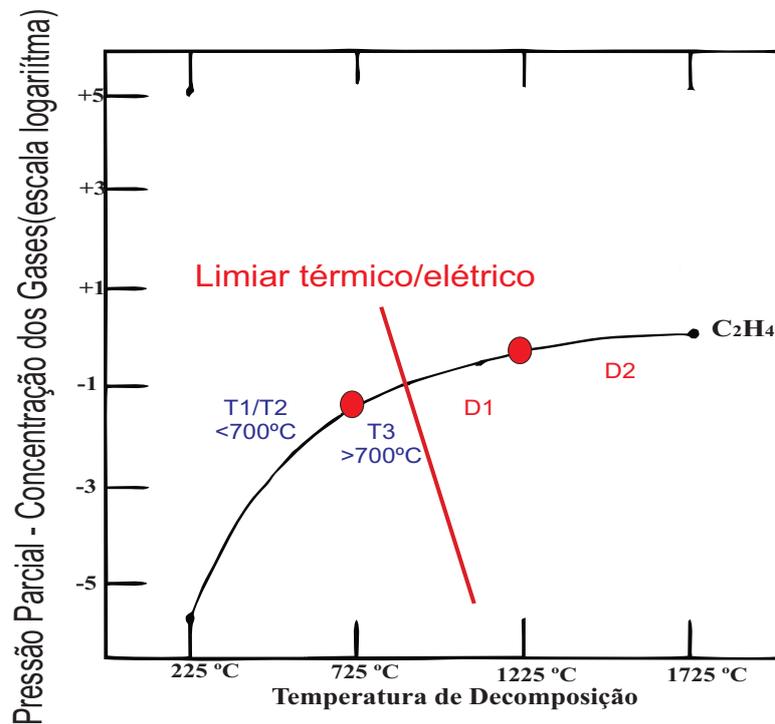
Fonte – PRÓPRIO AUTOR

Figura 37 – Rastreabilidade da porcentagem de etileno ( $PC_2H_4$ )



Fonte – PRÓPRIO AUTOR

Figura 38 – Divisão do etileno segundo a teoria de formação dos gases dissolvidos



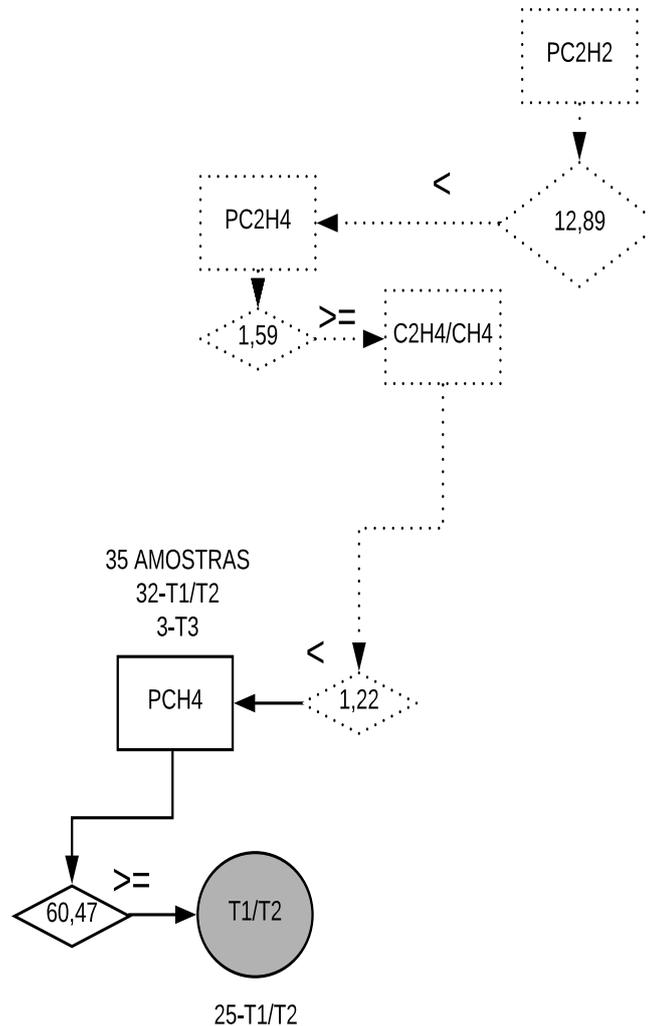
Fonte – PRÓPRIO AUTOR

a presença em grande quantidade de metano, conforme induzido pela árvore, indicando a presença do referido atributo para falhas de baixa temperatura. A razão entre o etileno e o metano também é evidenciada na Figura 40, pela qual a medida que a quantidade de  $C_2H_4$  aumenta, a de  $CH_4$  diminui, porém, para valores de temperaturas abaixo de  $700^\circ C$ , por menor que seja o valor de metano, este sempre será maior do que a quantidade de etileno, fazendo com que a relação esteja sempre com valores menores que 1, levando as amostras para a zona de falhas térmicas T1/T2.

- Hidrogênio( $H_2$ ): Atributo contido nos últimos níveis da DT. Pela análise da Figura 41, sua presença é dada pelo refinamento na busca pela classificação de amostras dentro do mesmo tipo de classe (elétricas ou térmicas) que atingem os níveis finais da árvore sem a devida separação. Para valores em torno de 525 ppm, classifica as amostras em T1/T2 e T3, e para valores em torno de 4010 ppm em falhas elétricas D1 e D2.

Comparando os resultados obtidos pela árvore de decisão com a teoria de formação dos gases (IEEE, 2008), observa-se (vide Figura 42) que o hidrogênio é o gás

Figura 39 – Rastreabilidade da porcentagem de metano ( $PCH_4$ )

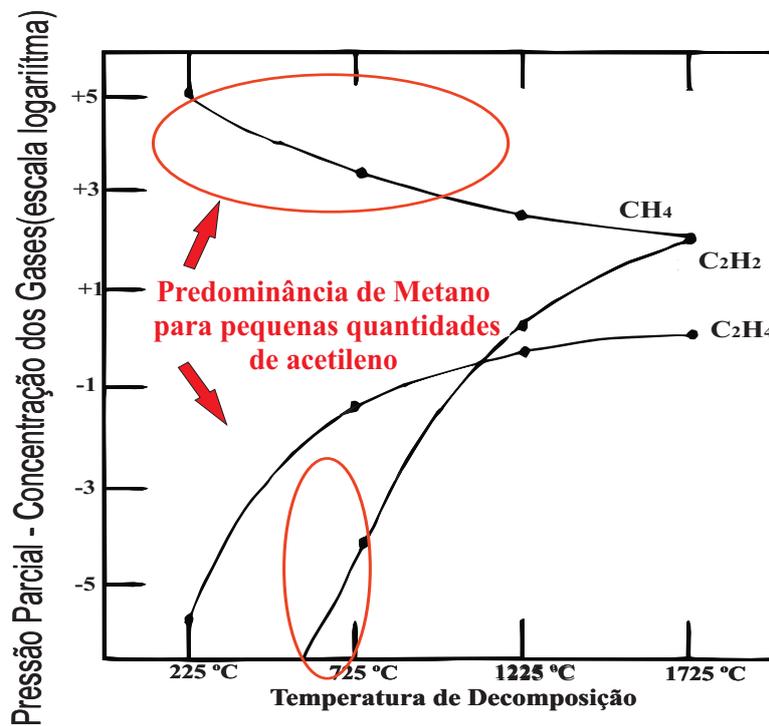


Fonte – PRÓPRIO AUTOR

com formação mais uniforme, tendendo a certa estabilidade a partir da temperatura de 725°C. Pela análise da referida figura, conclui-se ainda que existem duas regiões para a atuação do hidrogênio, em que a primeira região corresponde ao crescimento entre 225 e 725°C, na qual este atua, conforme indicado na árvore de decisão na separação de amostras térmicas e a segunda região corresponde a faixa estabilizada do atributo que atua na separação dentro do grupo das falhas elétricas.

Para análise de falhas de baixa energia, costuma-se utilizar também a relação ( $CH_4/H_2$ ), muito pelo fato da existência de um ponto de cruzamento entre esses dois gases. A análise da relação aponta que quando esta obtêm valores maiores que 1, a quantidade de metano se sobrepõe a de hidrogênio, estando a falta alo-

Figura 40 – Atributo metano segundo a teoria de formação dos gases dissolvidos

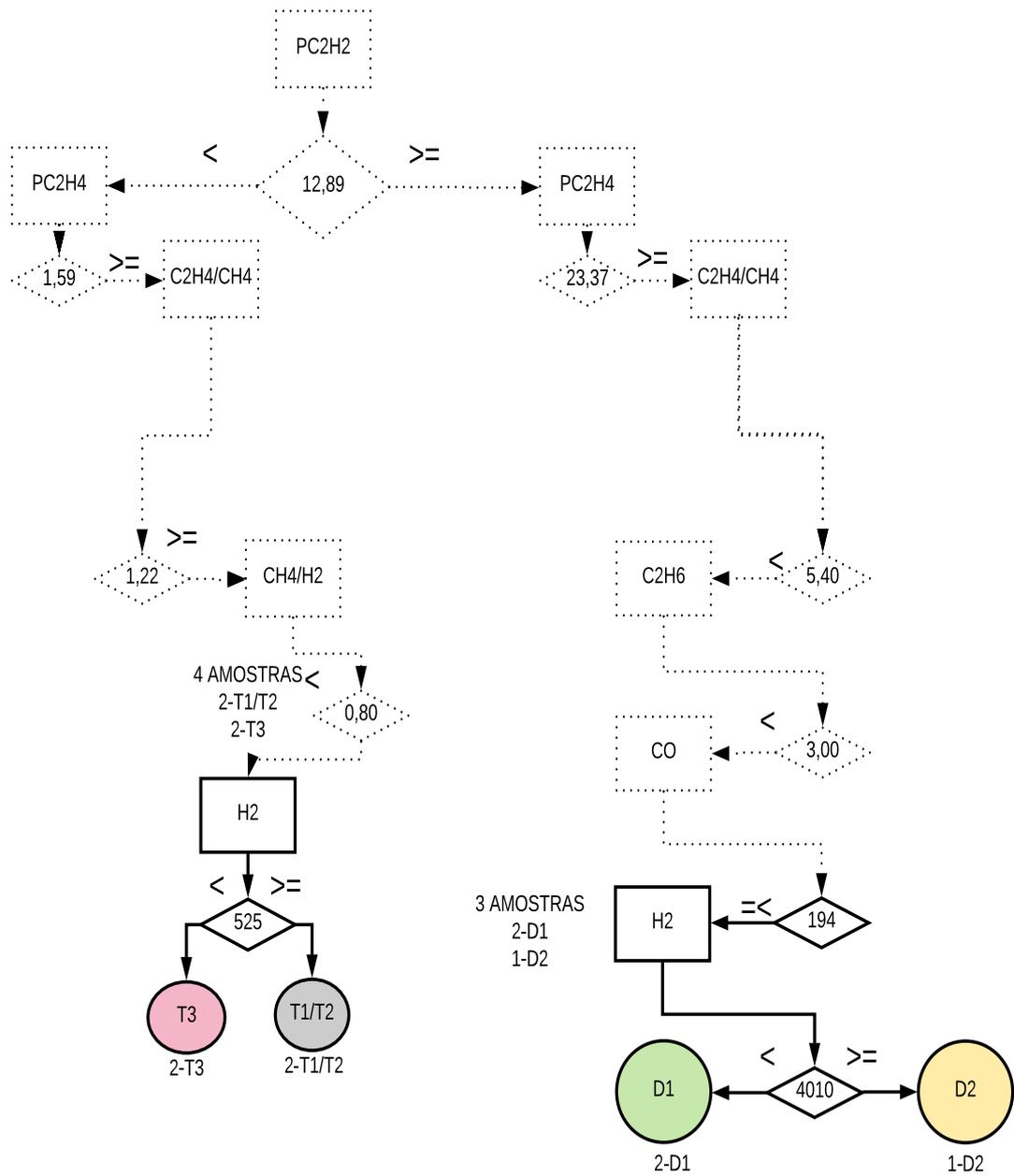


Fonte – PRÓPRIO AUTOR

cada na primeira região, correspondente a faltas térmicas. A Árvore de decisão final, ilustra essa característica, uma vez que associa tal relação ao ramo da árvore destinado a classificação desse tipo de falha.

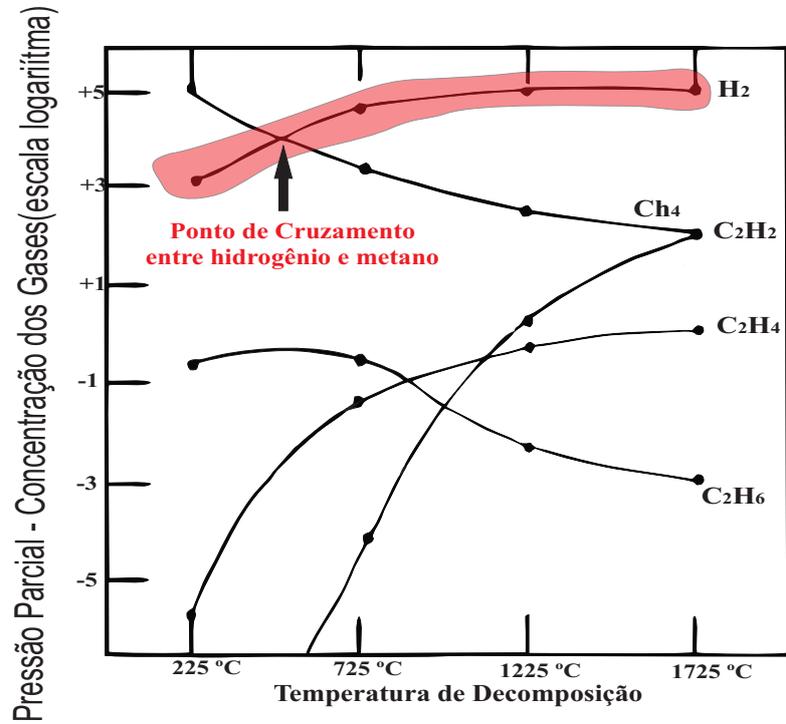
- Metano ( $CH_4$ ): Atributo presente no ramo da árvore cujo valor de Acetileno ( $\%C_2H_2$ ) é menor do que 12,89. Pela análise da Figura 43, infere-se que quando a amostra também contém valores baixos de  $\%C_2H_4$  (menores que 1,59), este atua diretamente na classificação de falhas do tipo PD, quando o seu valor supera 81,50. A Figura 40, ilustra a interação entre metano e acetileno para a classificação de falhas de baixa energia, como as descargas parciais, classificadas pela árvore de decisão, corroborando com a teoria de formação de gases.
- Monóxido de carbono ( $CO$ ): Atributo presente no ramo da árvore cujo valor de Acetileno ( $\%C_2H_2$ ) é maior do que 12,89. A Figura 44 apresenta que sua presença é dada pelo refinamento na busca pela classificação de amostras dentro da classe de falhas elétricas que atingem os níveis finais da árvore sem a devida separação. Com valores em torno de 194 ppm, classifica as amostras em D1 e

Figura 41 – Rastreabilidade do hidrogênio ( $H_2$ )



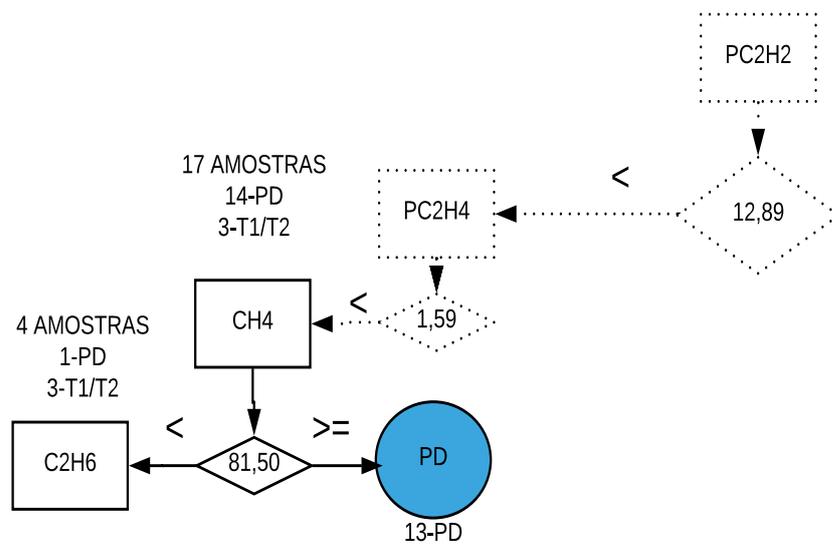
Fonte – PRÓPRIO AUTOR

Figura 42 – Atributo hidrogênio segundo a teoria de formação dos gases dissolvidos



Fonte – PRÓPRIO AUTOR

Figura 43 – Rastreabilidade do metano (CH<sub>4</sub>)



Fonte – PRÓPRIO AUTOR

D2.

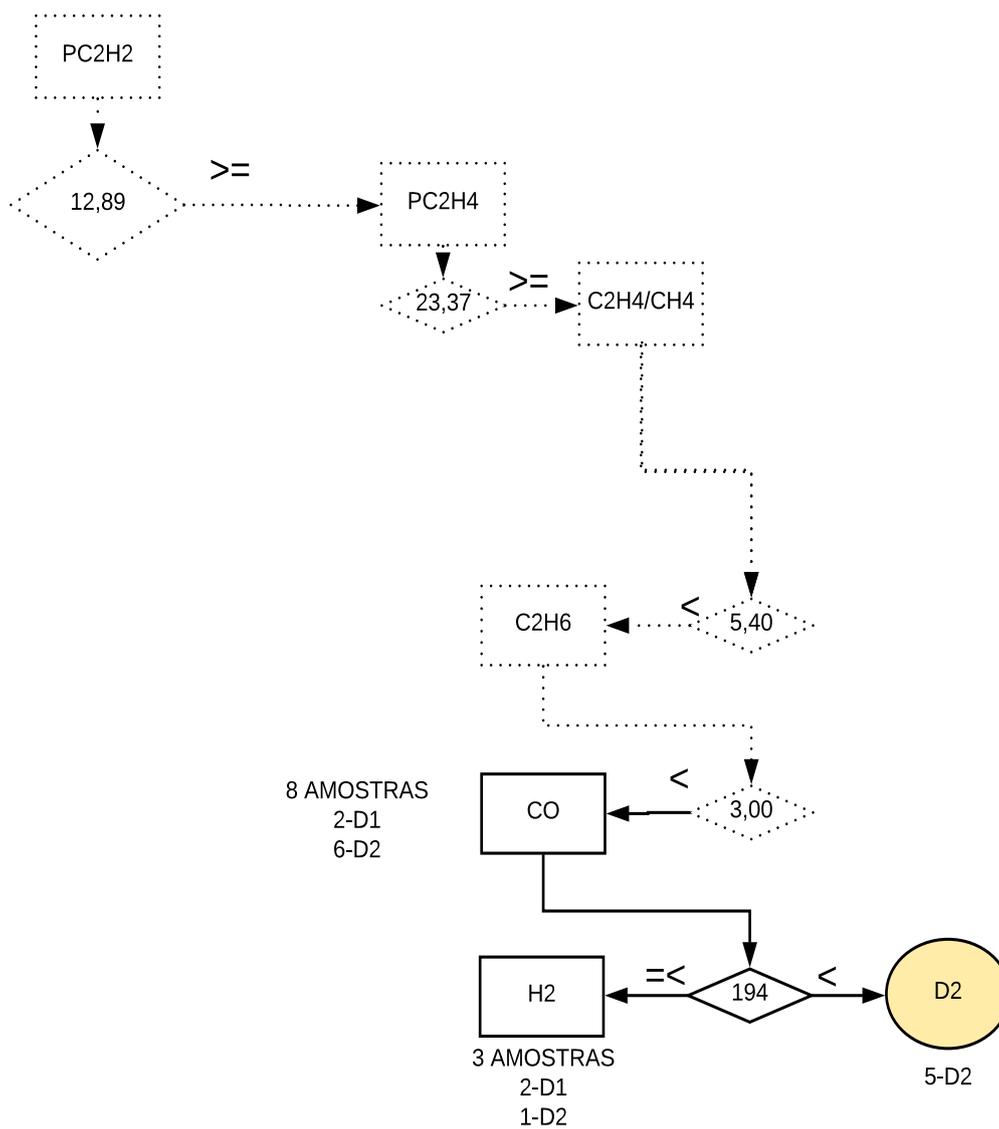
Segundo Barbosa (2013), Em uma situação da celulose sob temperatura extremamente alta, como sob arcos elétricos, a razão CO/CO<sub>2</sub> aproxima-se de 1:1, pois a geração de monóxido de carbono (CO) aumenta muito rapidamente em comparação à produção de dióxido de carbono (CO<sub>2</sub>). Porém, em uma situação de leve sobrecarga ou problemas de ventilação, onde ocorre leve sobreaquecimento, o CO<sub>2</sub> cresce muito mais rapidamente que o CO, portanto a taxa de CO/CO<sub>2</sub> fica na faixa de 1:20 a 1:10.

Analisando as amostras classificadas no ramo da árvore de decisão para valores de (%C<sub>2</sub>H<sub>2</sub>) menor que 12,89, observa-se que o valor de dióxido de carbono é superior ao de monóxido de carbono conforme mostrado na Figura 45, onde podemos inferir que a metodologia apresentada também apresenta conformidade no resultados encontrados para os óxidos provenientes do carbono que quantificam a degradação da celulose.

Ainda pela análise da Figura 45, podemos observar que a relação CO/CO<sub>2</sub> permanece próximo a zero para a maioria dos casos, corroborando os valores apresentados por Barbosa (2013) variando de 1:10 a 1:20.

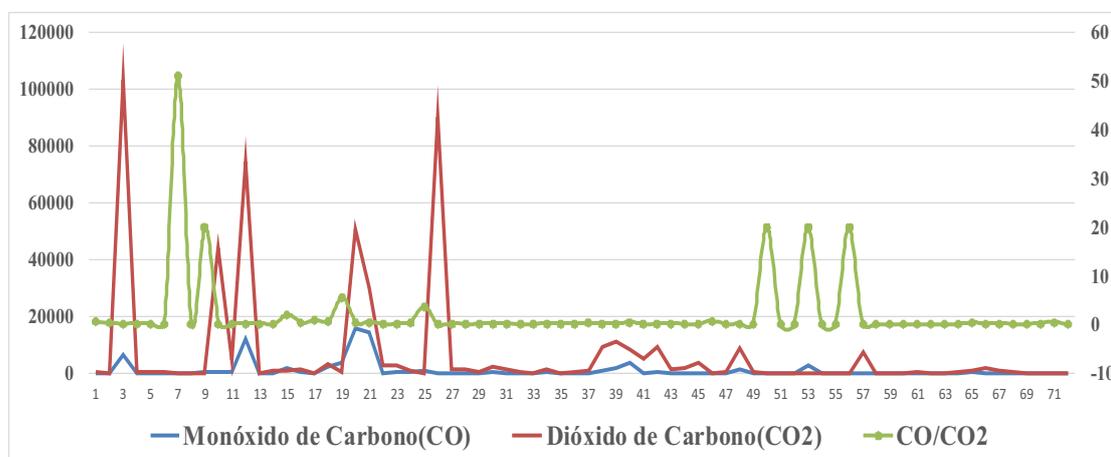
- Dióxido de carbono (CO<sub>2</sub>): Atributo utilizado para a classificação de falhas elétricas de baixa energia(D1) que fogem de valores convencionais(%C<sub>2</sub>H<sub>2</sub> maior que 18,89 e %C<sub>2</sub>H<sub>4</sub> menor que 23,37). Em concentrações maiores que 13 ppm, nos últimos níveis da DT(vide Figura 46), indica que a amostra trata-se de uma falha térmica T3. Tal classificação é condizente e relevante visto que, em ocorrências de temperaturas elevadas, internas ao transformador, existirá a degradação da celulose do papel isolante do enrolamento, tendo como subproduto da queima o depósito de sedimentos em forma de dióxido de carbono(CO<sub>2</sub>).
- Etano(C<sub>2</sub>H<sub>6</sub>):Pela análise da Figura 47, verifica-se que o referido atributo está presente nos dois ramos principais da DT(%C<sub>2</sub>H<sub>2</sub> maior e menor que 12,89). É um atributo importante, pois normalmente atua na separação das amostras térmicas das elétricas, remanescente das divisões principais, ocorridas nos primeiros níveis da árvore. Em valores altos(2345 ppm), profere a divisão de faltas D1 e T1/T2. Pela análise da Figura 48, observa-se que o etano também é um

Figura 44 – Rastreabilidade do monóxido de carbono(CO)



Fonte – PRÓPRIO AUTOR

Figura 45 – Atributos monóxido e dióxido de carbono para amostras classificadas como falhas térmicas e descargas parciais

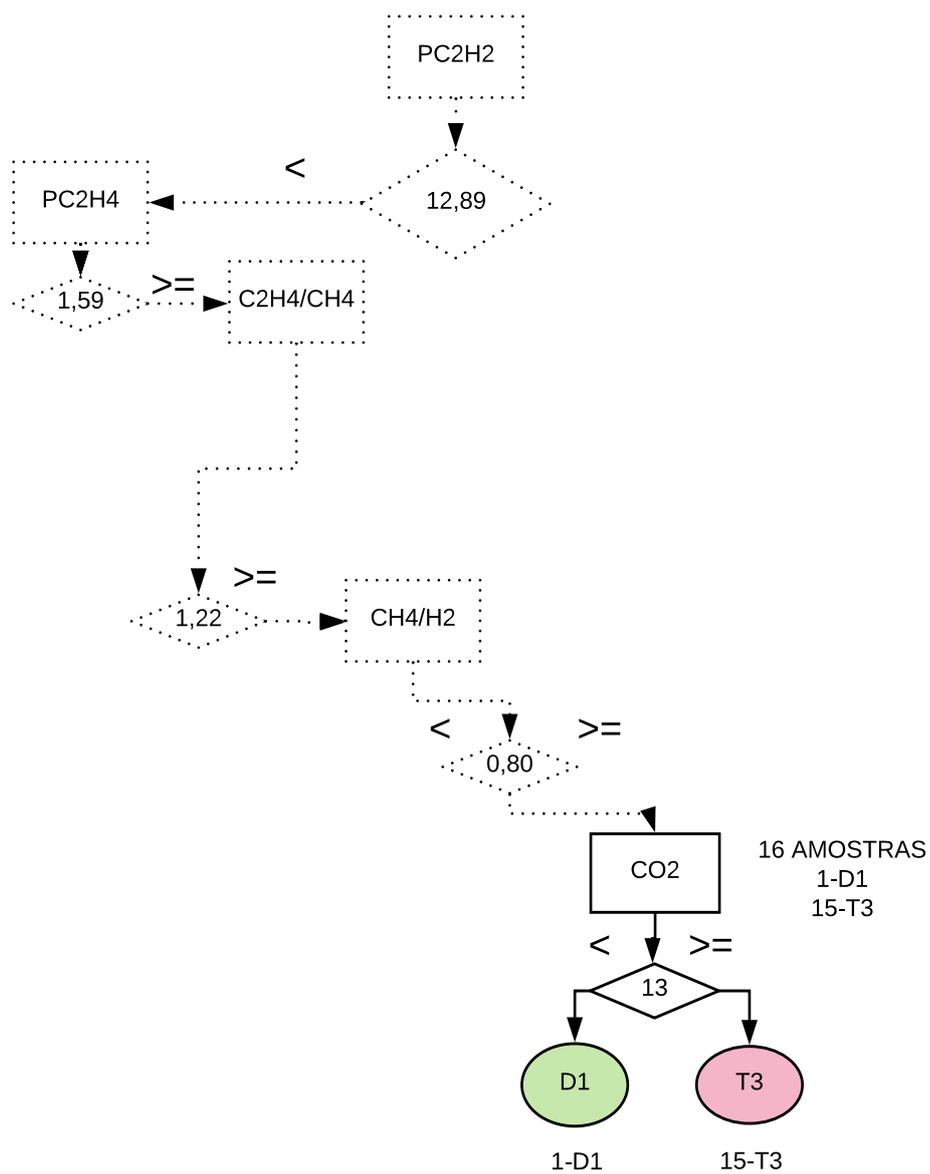


Fonte – PRÓPRIO AUTOR

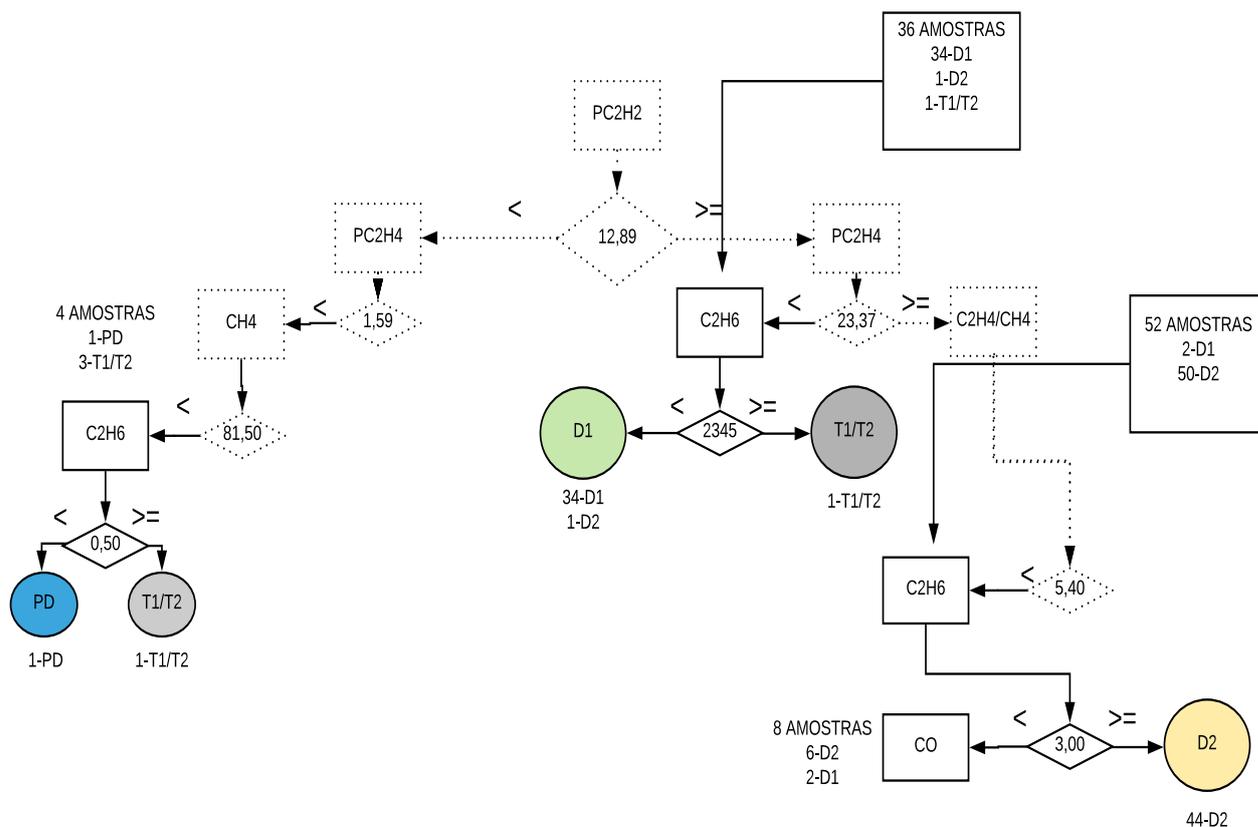
gás que se mantém presente em praticamente todo o espectro de temperatura, apresentando um comportamento crescente e gradual para baixas temperaturas, e decrescente para altas temperaturas, passando por uma região de inversão que se estende entre 700°C e 1250°C.

- Relação ( $C_2H_4/CH_4$ ): Presente nos dois ramos principais da DT ( $\%C_2H_2$  maior e menor que 12,89). Conforme indicado na Figura 49, atua nos níveis superiores da DT, sendo fundamental para classificações dentro da mesma classe de falha. Em se tratando de falhas térmicas, para valores em torno de 1,22, divide as faltas T1/T2 das T3. No outro ramo da árvore ( $\%C_2H_2$  maior que 12,89) no geral, divide as amostras D1 de D2. Ainda pela análise da Figura 48, observa-se que pela teoria de formação dos gases a relação decresce rapidamente no primeiro terço de temperatura e estabiliza-se no terço final
- Relação ( $CH_4/H_2$ ): Atributo presente no ramo da árvore cujo valor percentual de Acetileno ( $\%C_2H_2$ ) é menor do que 12,89. Em estágios intermediários da DT, conforme indicado na Figura 50, auxilia na divisão das amostras T1/T2 e T3. A análise da referida relação (vide Figura 51) configura-se como um dos pontos principais quando se utiliza a teoria da formação dos gases ratificando os

Figura 46 – Rastreabilidade do dióxido de carbono ( $CO_2$ )



Fonte – PRÓPRIO AUTOR

Figura 47 – Rastreabilidade do etano ( $C_2H_6$ )

Fonte – PRÓPRIO AUTOR

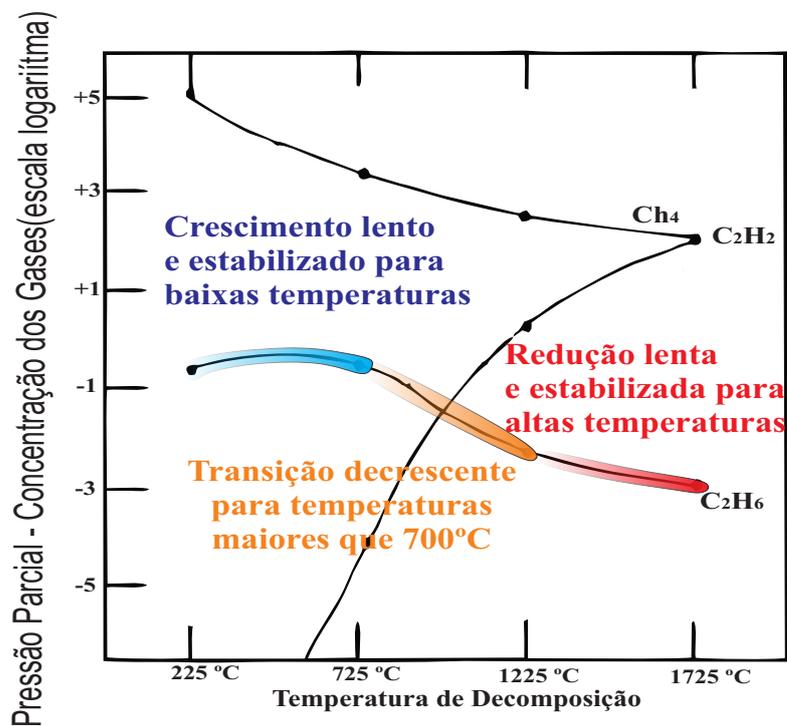
resultados apontados pela árvore de decisão.

O fato dos dois gases apresentarem crescimento inverso em todo o espectro de temperatura, definindo ainda um ponto de cruzamento no primeiro terço da escala é evidenciado na árvore, uma vez que a relação adquire valores menores que 1, classificando a amostra após o ponto de cruzamento e em determinado ramo, os valores para a relação atingem valores maiores do que a unidade, apontando para a região de baixas temperaturas localizada antes do ponto de cruzamento.

## 5.2 Regras Extraídas pela DT Proposta para Análise de Gases Dissolvidos no óleo de Transformadores de Força

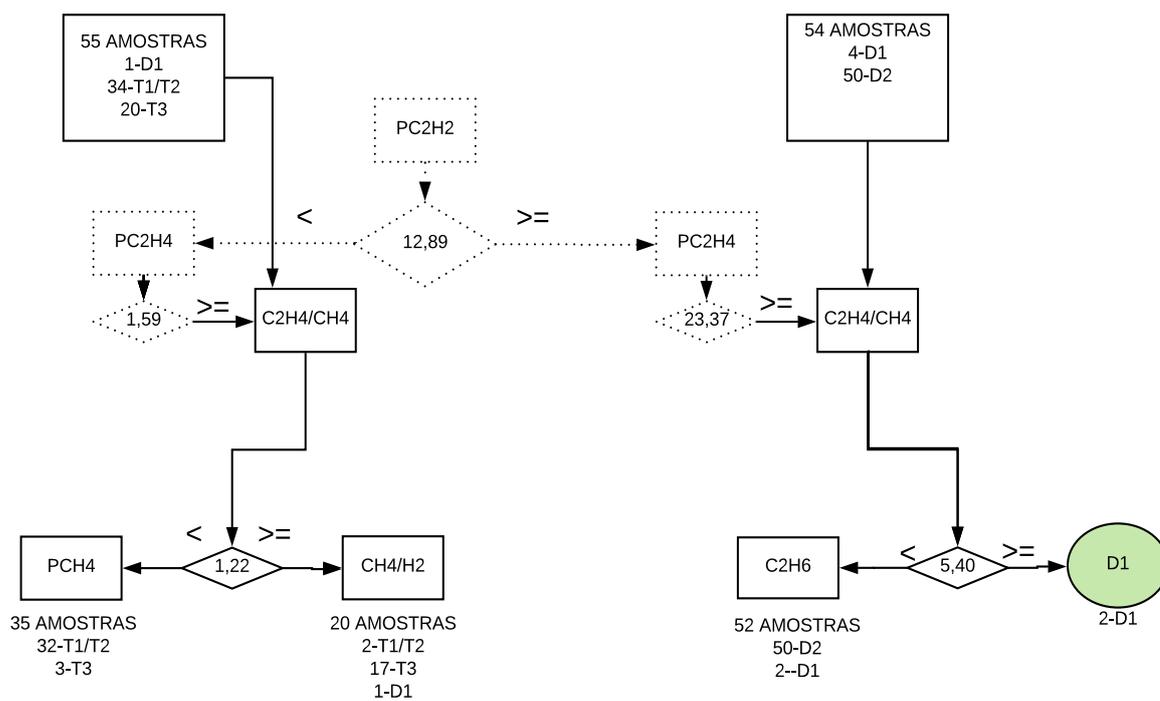
Após a indução da árvore de decisão final (seção 4.2) e a definição do novo espaço de entradas (seção 4.3), bem como a análise da aplicabilidade de cada um dos atributos escolhidos (seção 5.1), deve-se agora concluir o estudo objeto do presente trabalho com a definição das

Figura 48 – Atributo etano segundo a teoria de formação dos gases dissolvidos



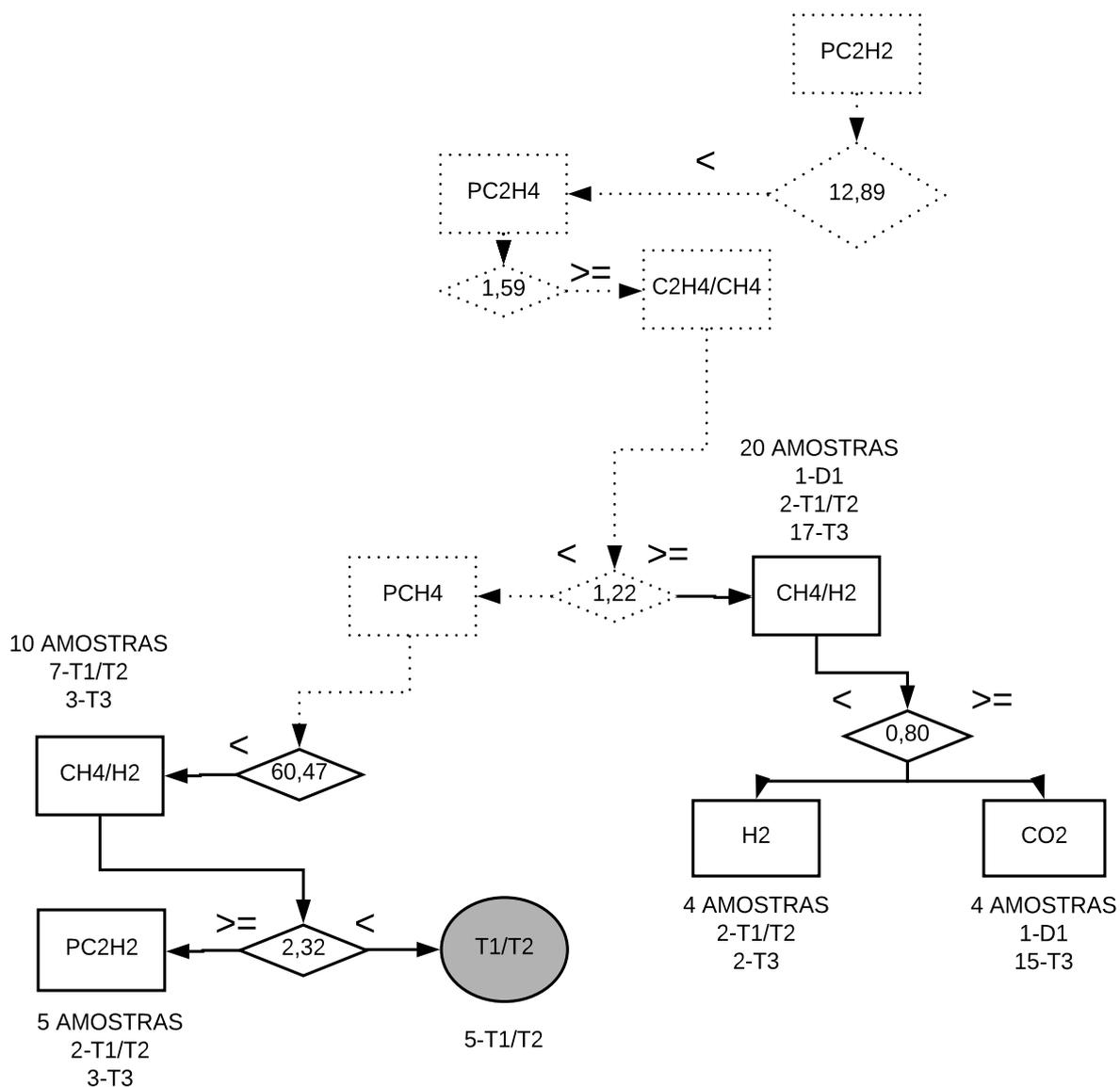
Fonte – PRÓPRIO AUTOR

Figura 49 – Rastreabilidade da relação (C<sub>2</sub>H<sub>4</sub>/CH<sub>4</sub>)



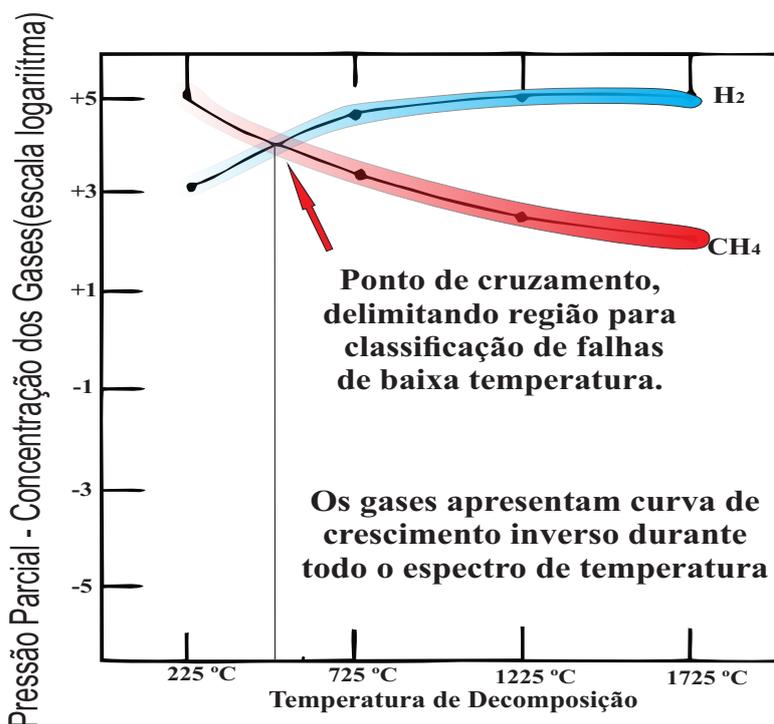
Fonte – PRÓPRIO AUTOR

Figura 50 – Rastreabilidade da relação( $CH_4/H_2$ )



Fonte – PRÓPRIO AUTOR

Figura 51 – Relação ( $CH_4/H_2$ ) segundo a teoria de formação dos gases dissolvidos



Fonte – PRÓPRIO AUTOR

regras extraídas a partir da análise da DT proposta.

As regras extraídas apresentadas na Tabela 14 serão o total de folhas presentes na DT(18), perfazendo o caminho da classificação de uma amostra qualquer, desde o nó raiz até o nó folha, alcançando no final a classificação da citada amostra pelo algoritmo.

A Tabela 15 mostra uma comparação entre o método desenvolvido neste trabalho com outros métodos utilizados em DGA para as 162 amostras em análise.

Evidencia-se pela Tabela 15, o grande potencial do método proposto em relação aos métodos já utilizados. A principal contribuição da metodologia proposta neste trabalho é a busca de informações que não são consideradas por métodos até agora implementados, a fim de otimizar a classificação de falhas em transformadores, evitando ao máximo, ocorrências de defeitos permanentes no equipamento.

### 5.3 Uma Nova Proposta para o Triângulo de Duval Utilizando Árvores de Decisão

O triângulo de Duval, como abordado anteriormente, configura-se como o método clássico mais bem aceito pela comunidade científica e aplicado em muitas concessionárias de energia para gestão da manutenção de transformadores de força. Duval, ao escolher os

Tabela 14 – Regras extraídas pela Árvore Proposta para Classificação DGA

Falta	%C2H2	% C2H4	% CH4	H2	CH4	CO	CO2	C2H6	C2H4/ CH4	CH4/ H2
PD	< 12.89	< 1.59			< 81.5			< 0.5		
PD	< 12.89	< 1.59			> 81.5					
T1/T2	< 12.89	< 1.59			< 81.5			> 0.5		
T1/T2	< 12.89	> 1.59	> 60.47						< 1.22	
T3	< 12.89 & < 1.42	> 1.59	< 60.47						< 1.22	> 2.32
T1/T2	< 12.89	> 1.59	< 60.47						< 1.22	< 2.32
T1/T2	< 12.89 & > 1.42	> 1.59	< 60.47						< 1.22	> 2.32
T3	< 12.89	> 1.59		< 525					> 1.22	< 0.8
T1/T2	< 12.89	> 1.59		> 525					> 1.22	< 0.8
D1	< 12.89	> 1.59					< 13		> 1.22	> 0.8
T3	< 12.89	> 1.59					> 13		> 1.22	> 0.8
D1	> 12.89	< 23.37						< 2345		
T1/T2	> 12.89	< 23.37						> 2345		
D1	> 12.89	> 23.37							> 5,40	
D2	> 12.89	> 23.37						> 3.0	< 5,40	
D2	> 12.89	> 23.37				< 194		< 3.0	< 5,40	
D2	> 12.89	> 23.37		> 4010		> 194		< 3.0	< 5,40	
D1	> 12.89	> 23.37		< 4010		> 194		< 3.0	< 5,40	

Tabela 15 – Percentual de Acerto na Classificação DGA Comparado a outros Métodos

Método	Acerto(%)
Metodologia Proposta	99,38%
Regras extraídas(AMORA,2013)	98,15%
Triângulo de Duval	88,03%
Dornenburg IEC C57.104 - 2008	63,25%
Rogers IEC C57.104 - 2008	56,41%

parâmetros metano, etileno e acetileno de forma percentual, torna balanceado a distribuição das amostras, facilitando o processo de classificação posterior.

Dessa forma, Duval consegue utilizando coordenadas triangulares, a classificação correta da maior parcela de falhas incipientes em transformadores utilizando apenas 3 tipos de gases combustíveis. Entretanto, a cada dia se torna mais necessário a otimização dos ativos e permanência em operação do equipamento durante todo o ciclo de vida útil, dessa forma, evidenciam-se as necessidades de busca de métodos que tendem cada vez mais o encontro da perfeição nos resultados.

Em 2012, Duval apresenta uma atualização do seu método, com a introdução de novos triângulos e 2 pentágonos, para aumento na precisão do tipo de falta envolvida, bem como a detecção de 5 subtipos de falhas detectáveis por DGA.

Tal fato, aumenta consideravelmente o percentual de acerto da técnica do triângulo,

porém, acaba com uma das maiores vantagens do método primário desenvolvido por Duval que é a simplicidade na análise.

A utilização de um triângulo seguido de outro, dependendo do resultado da análise do triângulo primário, torna a técnica trabalhosa e com mais variáveis para a análise da falha real. Já o uso do pentágono não é uma técnica de diagnóstico direto, tendo o especialista que ter uma maior sensibilidade para o entendimento do método proposto.

A utilização de árvores de decisão, nesse contexto, pode apresentar uma solução alternativa, ou seja, um novo formato para o triângulo proposto por Duval, otimizando as áreas de detecção das falhas, uma vez que ao algoritmo sejam apresentados os parâmetros do triângulo primário.

Com as regras extraídas da árvore proposta, agora com os mesmos parâmetros utilizados por Duval, pode-se traçar regiões em coordenadas triangulares, modificando as regiões originais a fim de aumentar o percentual de acerto do método.

As Tabelas 16 e 17 apresentam respectivamente as características associada à DT induzida, utilizando os parâmetros do triângulo de Duval e a matriz confusão para a referida classificação.

Tabela 16 – Características da DT Induzidas pelo Algoritmo C4.5 utilizando os parâmetros do Triângulo de Duval.

Árvore	Número de nós	Número de regras	% de acerto	Atributos apresentados
1	19	10	94,44%	%C <sub>2</sub> H <sub>2</sub> , %C <sub>2</sub> H <sub>4</sub> , %CH <sub>4</sub>

Tabela 17 – Matriz de Confusão para Classificação DGA Utilizando a Árvore Elaborada com os Parâmetros do Triângulo de Duval

	PD	D1	D2	T1/T2	T3
PD	14	0	0	0	0
D1	0	36	2	0	1
D2	0	1	50	0	0
T1/T2	3	0	0	33	2
T3	0	0	0	0	20

Pela análise da Tabela 16, verifica-se que a utilização de árvores de decisão, otimiza o método de três gases de Duval, elevando a acurácia de 88,03% para 94,44% utilizando a mesma

base de dados.

A matriz de confusão (vide Tabela 17) apresenta os pontos de classificação incertos do método proposto. Durante o processo, a ocorrência de erros pontuais é presente em quase todas as classes, porém em um número bem menor do que ao usar o triângulo convencional.

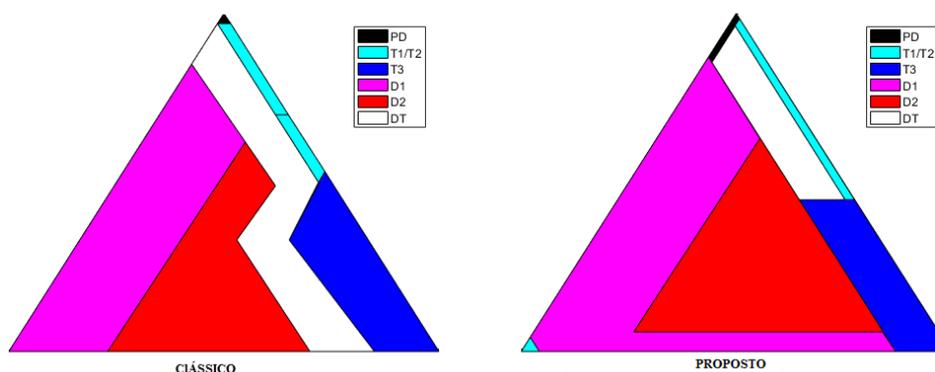
As regras extraídas da DT induzida utilizando os parâmetros de Duval são apresentadas na Tabela 18.

Tabela 18 – Regras extraídas pela Árvore Proposta para Classificação DGA

Falta	% C2H2	% C2H4	% CH4
PD	< 12.89	< 1.59	
T3	< 12.89	> 1.59	< 44.89
T1/T2	< 12.89	> 1.59	> 44.89 & > 60.47
T3	< 12.89 & > 0.30 & < 1.42	> 1.59	> 44.89 & < 60.47
T1/T2	< 12.89 & > 0.30 & > 1.42	> 1.59	> 44.89 & < 60.47
T1/T2	< 12.89 & < 0.30	> 1.59	> 44.89 & < 60.47
D1	> 12.89 & < 95.77	< 23.77	
T1/T2	> 12.89 & > 95.77	< 23.77	
D2	> 12.89	> 23.77	> 5.81
D1	> 12.89	> 23.77	< 5.81

Traçando as regiões correspondentes a cada tipo de falta mapeada pela DT em análise, pode-se propor uma nova configuração para o triângulo de Duval em detrimento ao convencional, ambos exibidos na Figura 52.

Figura 52 – Triângulo de Duval Clássico x Triângulo Proposto.



Fonte – PRÓPRIO AUTOR

Pela análise do triângulo proposto a partir das regras extraídas pela DT proposta com

os mesmos parâmetros do triângulo de Duval observa-se que a região definida para Descargas Parciais (PD) se alonga um pouco mais no sentido decrescente de  $\%CH_4$  e crescente de  $\%C_2H_2$ .

Essa alteração, confirmada agora pelo uso de DT, já tinha sido mapeada por Duval na concepção do triângulo 5, utilizado em transformadores com avarias de altas temperaturas e óleo mineral. Este é utilizado quando no triângulo clássico as faltas são mapeadas como faltas de alta temperatura. Na ocasião, utiliza-se o novo triângulo (cinco) para busca de subtipos de falhas, sendo neste, a área mapeada para PD, a mesma encontrada pelo triângulo proposto.

O alongamento da região originalmente mapeada como D1 para a zona de incidência de qualquer valor de  $\%C_2H_2$ , desde que permanecendo com baixos índices de  $\%CH_4$  é similar a região do triângulo 5 mapeada como sobreaquecimento para temperaturas  $T < 250^\circ C$ .

A região definida como D2 (Falhas de alta energia), obteve um aumento de sua abrangência pelo mapeamento das regras da Árvore de Decisão, bem como a zona de DT (falhas térmicas associadas a elétricas) apresentou retração.

Observa-se ainda que para o triângulo proposto, a zona de falhas térmicas abaixo de  $300^\circ C$  (T1) e falhas térmicas entre  $300^\circ C$  e  $700^\circ C$  (T2), foi unificada. Esse fato é explicado pois no tratamento do banco de dados, não foi feita diferenciação entre amostras T1 e T2, sendo ambas alocadas na mesma classificação (T1/T2).

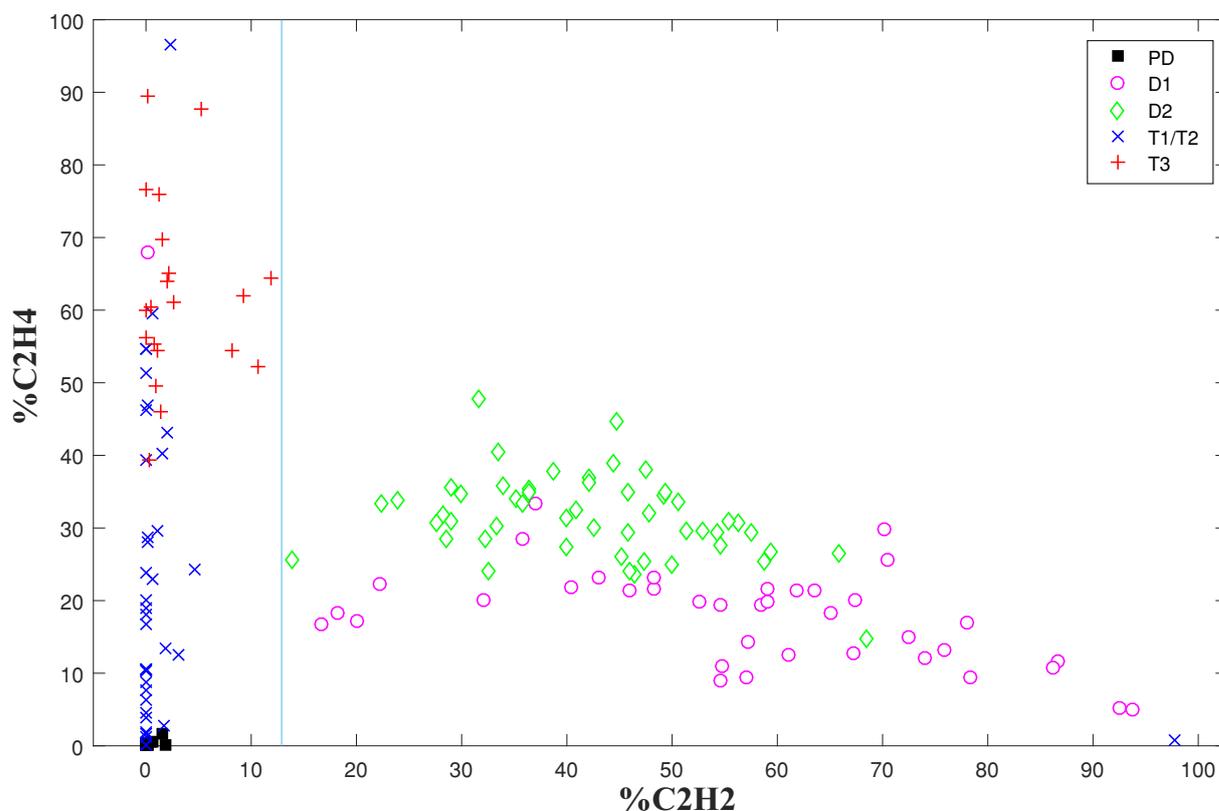
Por fim, o triângulo proposto apresenta uma nova zona, correspondente a falhas T1/T2. Essa zona é caracterizada por amostras que contém um alto valor de  $\%C_2H_2$ , acompanhado de valores muito baixos de  $\%CH_4$ .

#### 5.4 A Regra do Acetileno para classificação DGA

O percentual relativo de acetileno ( $\%C_2H_2$ ), demonstrou no desenvolvimento do presente trabalho, papel fundamental para o processo de classificação DGA, como sendo o atributo que em qualquer situação divide as amostras da melhor forma.

Não distante desse fato, o percentual relativo de etileno ( $\%C_2H_4$ ) tem acompanhado a tendência do  $\%C_2H_2$  como atributo balanceado, auxiliar na divisão principal.

A Figura 53 ilustra as 162 amostras trabalhadas dispersas no plano cartesiano formado pelos percentuais relativos citados. Pela análise da referida figura, observa-se que todas as amostras PD, T1/T2 e T3, com exceção de uma amostra T1/T2, totalmente deslocada do centróide de sua classe (podendo representar assim um *outlier*), independente do valor de  $\%C_2H_4$ , estão abaixo do valor 12,89 para  $\%C_2H_2$ .

Figura 53 – %C<sub>2</sub>H<sub>2</sub> x %C<sub>2</sub>H<sub>4</sub>

Fonte – PRÓPRIO AUTOR

Verifica-se também que após o valor de 12,89 de %C<sub>2</sub>H<sub>2</sub>, estão localizadas todas as amostras D1 e D2, com exceção de uma amostra D1, também totalmente deslocada em relação ao centróide da sua classe (podendo também representar um *outlier*).

Dessa forma, pode-se propor uma regra para classificação imediata de falhas em transformadores, divididas em 02 grupos que refletem a dimensão inicial do tipo de falta em questão.

Tal regra propõe que para valores de %C<sub>2</sub>H<sub>2</sub> abaixo de 12,89, tem-se a ocorrência de falhas térmicas(T1/T2 ou T3) ou descargas parciais(PD), e que para valores de %C<sub>2</sub>H<sub>2</sub> maiores que 12,89, tem-se a ocorrência de falhas elétricas, sejam de baixa(D1) ou alta energia(D2).

Pela regra proposta, comparada ao triângulo de Duval, existe a divisão das falhas do tipo DT (térmicas e elétricas), pelo valor de 12,89, sendo as de baixo valor (menor que o limiar de separação), associadas a falhas térmicas e as de valor superior, associadas a falhas elétricas, obedecendo assim a divisão proposta.

Pelo triângulo proposto, as falhas do tipo DT, com valores de %C<sub>2</sub>H<sub>2</sub> maiores que

12,89, são classificadas como D2, e as menores do que o limiar, continuam sendo classificadas como DT.

## 5.5 Utilização da DT Proposta e da Regra do Acetileno em DGA

Afim de demonstrar a aplicabilidade e conformidade de dois dos métodos propostos pelo trabalho, buscou-se a análise de amostras cromatográficas de transformadores de força abordados em artigos científicos e alguns casos de transformadores operando em concessionárias de energia.

### 5.5.1 Casos encontrados em Siada e Hmood (2013)

Em Siada e Hmood (2013), foi proposto um método utilizando lógica fuzzy para identificar problemas em transformadores de força utilizando DGA. Para o presente trabalho, são comparados os resultados de saída do modelo fuzzificado proposto com os seguintes parâmetros:

- *F1*: Falha térmica na celulose ou no óleo;
- *F2*: Descarga elétrica de baixa energia (Corona);
- *F3*: Falha elétrica (Arco);

A Tabela 19, detalha quatro amostras utilizadas no citado trabalho, a classificação pelo método fuzzificado e pela DT proposta.

Tabela 19 – Comparação na classificação: Método Fuzzificado x DT Proposta

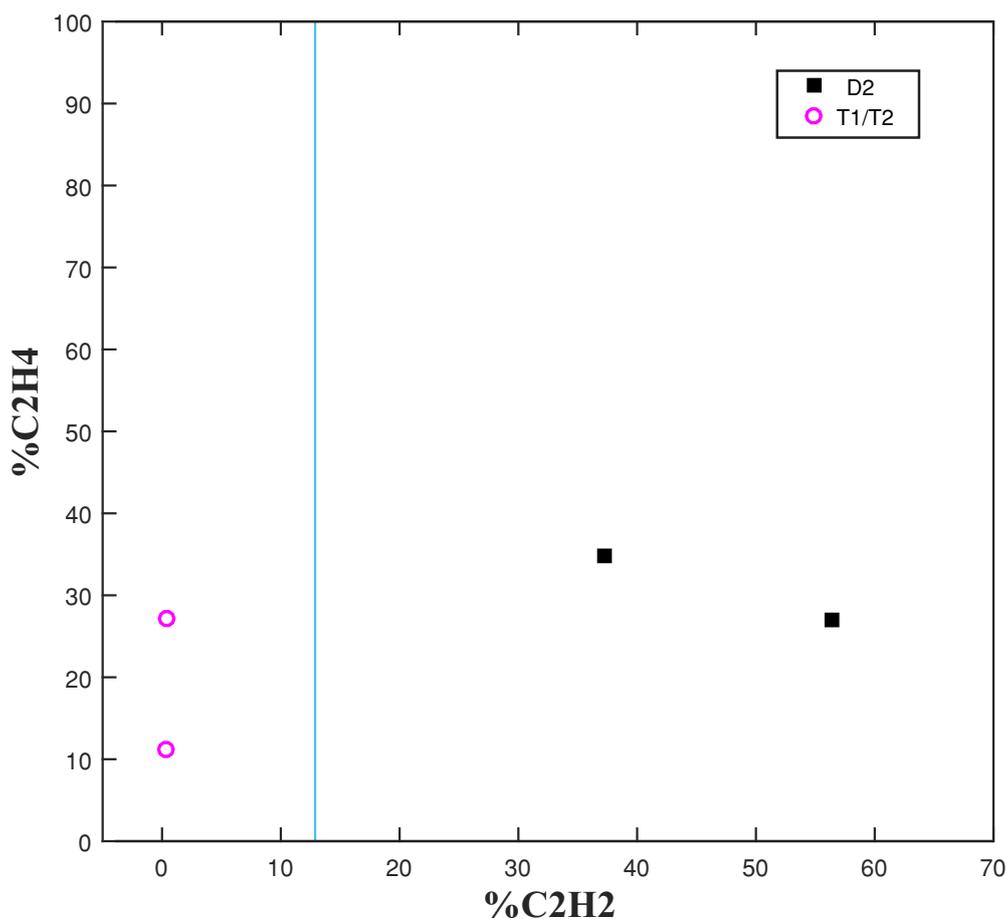
Amostra	H2	CH4	C2H2	C2H4	C2H6	CO	CO2	SIADA	DT PROPOSTA
1	124	166	0	59	87	530	3750	F1	T1/T2
2	103	74	0	9	80	754	2605	F1	T1/T2
3	441	207	261	224	43	161	1123	F3	D2
4	217	286	884	458	14	176	1544	F3	D2

Pela análise da tabela acima, denota-se que a classificação das amostras pela DT proposta foi a mesma efetuada pelo modelo fuzzificado, com o incremento de direcionar dentro do conjunto possível de falhas térmicas e elétricas, o tipo de falha que provavelmente esteja ocorrendo.

No modelo fuzzificado, a classificação restringe-se em direcionar a falha para térmica. Já na DT proposta a classificação aponta para uma falha T1/T2, restringindo o subconjunto solução, procedimento idêntico ao ocorrido na classificação F3 e D2.

Utilizando a Regra do acetileno, tem-se a disposição abaixo das amostras (vide Figura 54).

Figura 54 – Regra do Acetileno para os quatro casos em Siada e Hmood, (2013)



Fonte – PRÓPRIO AUTOR

Observa-se pela Figura 54 que a classificação das amostras nos dois métodos abordados obedecem a regra do acetileno proposta no presente trabalho.

### 5.5.2 Casos encontrados em Wang (2014)

Em Wang, (2014), o autor utiliza a correlação e análise de séries temporais para realizar a análise da deterioração sólida do papel em transformadores de força com base em DGA.

Para a realização do estudo acima, o autor utilizou amostras de dois transformadores, sendo que um apresentava falha térmica e outro falha elétrica, e através do acompanhamento da evolução da falta inferir acerca da evolução desta e diagnóstico do transformador.

A Tabela 20, indica o acompanhamento das amostras para o caso do transformador em falha térmica, bem como o diagnóstico pela DT Proposta.

Tabela 20 – Diagnóstico do Acompanhamento do transformador em falha térmica pela DT proposta

Data da coleta	H2	CH4	C2H2	C2H4	C2H6	CO	CO2	DT Proposta
18/09/2001	5.7	3.3	0	2.6	1.8	611.6	3671.5	T1/T2
12/10/2001	9.7	3.4	0	4.1	1.2	705.2	3803.3	T1/T2
7/11/2001	20.6	63.3	0	91.3	20.9	614.3	3863.2	T3
12/11/2001	22.3	53.1	0	86.5	20.7	517.2	15477	T3
27/11/2001	32	70.5	0	78	26.9	452.9	2789.6	T1/T2
12/12/2001	37.3	211.6	0	170.4	43.2	666.5	2709.1	T3
30/12/2001	31.6	222.6	0	184.4	48.3	615.4	2536.9	T3
07/01/2002	31	246.4	0	224.4	48.7	543.2	2260.4	T3
15/01/2002	45.4	212.2	0	170.7	46.3	739	3586.9	T3
24/01/2002	70.2	123.1	0	179.5	85.2	638.7	3527.5	T3
10/02/2001	61.3	386.6	0	357.2	85.7	828.5	3864.4	T3
19/02/2002	72	271.7	0	281.7	95.3	679.4	6350	T3
26/02/2002	87.4	310.6	0	306.2	95.4	675.1	3595.3	T3
02/03/2002	66.8	328.7	0	353.8	100.1	712.9	3640.5	T3
10/03/2002	76.4	385.7	0	436.9	88.6	880.5	4800.6	T3
11/03/2002	277	350.6	0	409	113.5	892.4	4823.1	T1/T2
13/03/2002	190.9	410.2	0	435	121.2	989.9	4979.4	T1/T2
14/03/2002	174.8	460.4	0	410.8	114.2	847.7	4861.6	T3

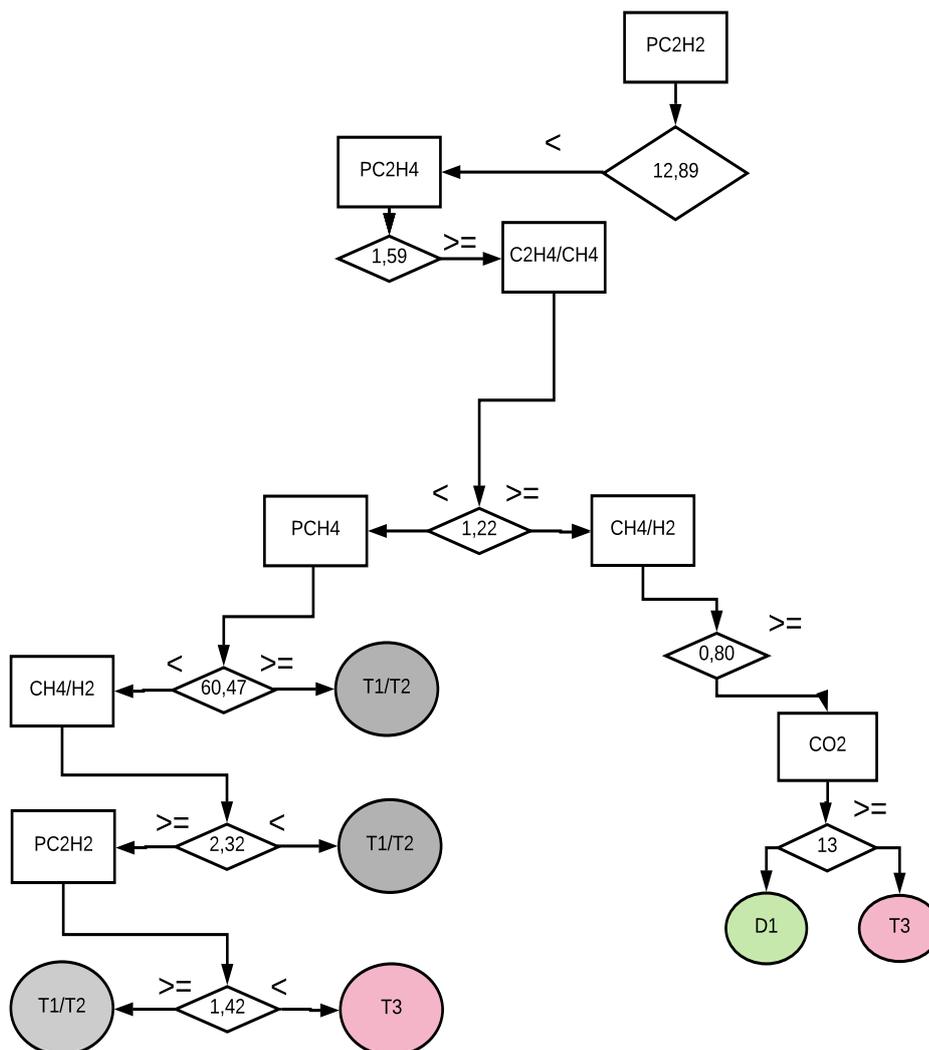
Pela análise da Tabela 20, observa-se que a DT proposta pelo presente trabalho pode também ser utilizada para o acompanhamento da evolução da falha incipiente em transformadores, pois para o transformador permanentemente em falha térmica, para todos os casos a DT indica falha térmica, seja T1/T2 ou T3.

A Figura 55 ilustra os ramos percorridos na DT para a classificação das amostras citadas, desde a raiz até a classificação nas folhas.

Para a classificação acima, observou-se que a divisão das amostras em T1/T2 ou T3 ocorreu no nó  $CH_4/H_2$ , no qual para amostras que contém esse atributo menor que 2,32 foram classificados como T1/T2 e maiores do que este valor em T3.

Observou-se também que em apenas 01 caso, a amostra continha o valor da relação

Figura 55 – Ramos da DT proposta utilizados na classificação das amostras em Wang (2014)



Fonte – PRÓPRIO AUTOR

$C_2H_4/CH_4$  maior que 1,32, porém, sendo classificado como T3 ao final do ramo consequente.

A Tabela 21 apresenta o acompanhamento da coleta e análise das amostras pela DT proposta para o caso do transformador em falha elétrica.

Tabela 21 – Diagnóstico do acompanhamento do transformador em falha elétrica pela DT proposta

Data da coleta	H2	CH4	C2H2	C2H4	C2H6	CO	CO2	DT Proposta
20/10/2011	31.33	10.52	6.09	4.01	1.98	275.25	3722.59	D1
28/10/2011	32.77	10.22	6.11	3.63	1.52	260.18	1677.22	D1
3/11/2011	41.36	11.08	6.21	4.01	1.75	312.12	1715.23	D1
10/11/2011	36.4	11.29	6.03	3.84	1.79	281.51	1633.58	D1
18/11/2011	43.8	10.98	6.1	3.85	1.81	377.6	1685.3	D1
24/11/2011	34.24	10.1	6.19	3.8	1.54	253.18	1691.91	D1
1/12/2011	39.94	11.77	6.38	4.36	1.88	293.24	1742.11	D1
5/12/2011	38.1	11.56	7.31	4.75	1.92	277.54	1693.72	D1
12/12/2011	35.5	10.73	7.23	4.32	1.94	262.76	1679.69	D1
19/12/2011	35.13	11.27	7.22	4.48	1.87	257.11	2699.09	D1
23/12/2011	36.81	10.82	6.88	4.3	1.76	259.7	2690.08	D1
21/02/2012	35.47	11.41	5.87	4.11	1.89	259.75	1588.82	D1
19/03/2012	33.82	10.9	7.2	4.4	1.9	258.06	1548.5	D1
21/03/2012	32.9	10.9	7.61	5.09	2.33	277.43	1551.9	D1

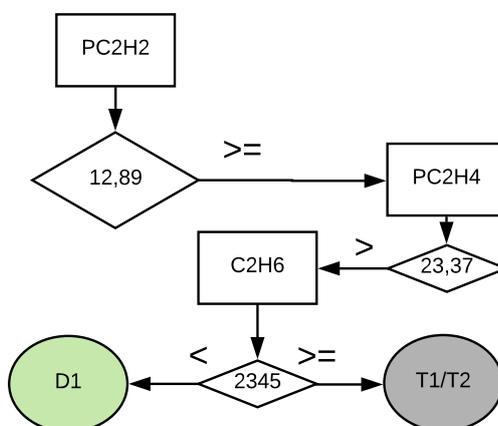
Observando a Tabela 21, verifica-se que todas as amostras foram classificadas como D1 pela DT proposta. Para os casos citados, existe a presença do acetileno porém em quantidades minoritárias, classificando assim as falhas em elétricas, porém em um grau primário de evolução (D1).

Pela regra do Acetileno, todas as amostras tem o atributo  $\%C_2H_2$ , maior que 12,89, configurando-se assim como falhas elétricas.

A Figura 56 ilustra os ramos percorridos na DT para a classificação das amostras citadas, desde a raiz até a classificação nas folhas.

Pela análise da Figura 56, verifica-se a facilidade com que a DT proposta classifica as amostras em D1. Uma vez que todas tem valores do atributo  $\%C_2H_4$  menores que 23,37 e  $C_2H_6$  inferiores a 2345, o ramo indicado é utilizado para classificação geral dos casos apresentados.

Figura 56 – Ramos da DT proposta Utilizados na classificação das amostras de falhas elétricas em (WANG,2014)



Fonte – PRÓPRIO AUTOR

## 5.6 Considerações Parciais

Após a indução da árvore de decisão final para melhor classificação DGA (Capítulo 4), na seção 5.1 foi efetuado um minucioso estudo em relação a rastreabilidade de cada um dos atributos escolhidos pelo algoritmo C4.5 para compor a DT proposta.

Na referida seção foi abordado um a um os dez atributos, evidenciando a utilização de cada um destes na classificação de determinada falta. Na seção 5.2 foram apresentadas as regras extraídas da DT proposta, utilizadas para classificação de falhas incipientes em transformadores. As regras são do tipo SE-ENTÃO, de fácil interpretação e aplicação.

Na seção 5.3 apresentou-se uma nova proposta para o Triângulo de Duval, método clássico de melhor aceitação para análise DGA, com base em regras extraídas da DT proposta, quando a esta, são apresentados os mesmos atributos pertencentes utilizados por Duval, obtendo assim uma assertividade considerável quando comparada ao método clássico.

Na seção 5.4 apresentou-se ainda a Regra intitulada por *Regra do Acetileno*, que com base no valor de 12,89 para a porcentagem relativa de acetileno, divide as amostras cromatográficas em térmicas e elétricas.

Por fim, na seção 5.5 utilizou-se alguns casos recentemente estudados e apresentados em artigos científicos para aplicação e comparação dos métodos desenvolvidos e propostos durante o desenvolvimento do presente trabalho.

No próximo capítulo, serão apresentadas as conclusões gerais acerca do presente trabalho, como também os trabalhos futuros a serem desenvolvidos.

## 6 CONCLUSÕES

Métodos tradicionais de análise DGA são válidos pois conseguem de forma simples e direta o diagnóstico da maior parte de amostras, empregando quantidade reduzidas de regras para tal atribuição. O método de Duval nesse contexto configura-se como o de melhor aproveitamento entre os métodos clássicos por utilizar as concentrações relativas dos gases que juntos abrangem boa parte dos casos de falhas incipientes em transformadores.

Com a disseminação e aplicação das técnicas de IC para análise de gases dissolvidos em transformadores, obteve-se um ganho significativo no percentual de acerto nas classificações, visto que esses métodos trabalham de forma satisfatória, situações de complexidade computacional considerável.

Nesse contexto, o uso de Árvores de decisão como método alternativo para análise DGA apresenta elevada eficiência, visto que trabalha de forma simples com atributos descon siderados por métodos tradicionais. O algoritmo C4.5 por utilizar como métrica para seleção de atributos a razão do ganho, consegue extrair ao máximo as informações disponibilizadas por cada classe, otimizando desta forma a classificação de amostras, anteriormente classificadas como erradas por métodos tradicionais.

Como resultado, o algoritmo C4.5 utilizado na indução de árvores de decisão, sempre apresentou uma alta performance na classificação de falhas incipientes, inicialmente com 98,15% de acurácia, e 99,38% para o caso final considerado.

Com a análise da árvore de decisão proposta pelo trabalho, foi possível mapear a aplicação dos atributos na classificação de falhas, direcionando de forma rápida e segura para o tipo de anomalia em ocorrência tão somente analisando os atributos a esta relacionada.

A extração de regras disjuntivas e sequenciais a partir da DT proposta, infere a possibilidade de com uma simples implementação computacional, obter-se o diagnóstico do tipo de falha a partir de dados cromatográficos com um elevado acerto. Tal fato sugere ainda que os tipos de falhas mapeadas, independente do seu grau de severidade dentro do subconjunto a qual esta pertence, seguem um padrão pré-definido e mapeado em um dos ramos disponíveis na árvore de decisão proposta.

Com a nova proposta para o triângulo de Duval, sugerida pelo presente trabalho, foi possível unir a simplicidade do triângulo idealizado com a eficiência do algoritmo de DT, evoluindo na classificação das amostras de 88,03% para 94,44%.

O presente trabalho também propôs uma regra para o gás acetileno. Tal regra, para

valores menores que 12,89 (concentração relativa de acetileno), classifica os tipos de falhas em transformadores em DP, T1/T2 E T3. Para valores acima deste, as falhas são classificadas em D1 e D2, sugerindo assim um parâmetro divisor de falhas elétricas e térmicas.

Por fim, conclui-se que a utilização da técnica de árvores de decisão para diagnóstico DGA, constitui uma ferramenta importante e promissora, seja na análise propriamente dita como em conjunto com técnicas tradicionais, estabelecendo assim de forma rápida e segura a ocorrência precisa da falha interna ao transformador, permitindo que ações corretivas e eficazes venham a ser tomadas, evitando assim a queima permanente e inutilização do equipamento.

## 6.1 Trabalhos Futuros

Como sugestão para trabalhos a serem desenvolvidos, motivados pelos assuntos e conclusões abordadas na presente Dissertação, tem-se a utilização do algoritmo *Random Forest* para a classificação DGA. Tal algoritmo embora pouco explorado na literatura, devido ao seu caráter "caixa preta" interno ao algoritmo, pode inferir em resultados satisfatórios e similares aos encontrados com a utilização de árvores de decisão, por se tratar de uma evolução desta.

Outro trabalho a ser abordado é o desenvolvimento das técnicas do novo triângulo de Duval e Regra do acetileno, utilizando outros casos existentes na literatura DGA e casos de transformadores em operação, bem como o desenvolvimento de uma solução para diagnóstico DGA, embarcada em circuito microprocessado e conectada diretamente a equipamentos coletores e amostradores de teor de gases em transformadores para diagnóstico rápido, eficiente e local de anomalias em transformadores de força.

Por fim, a aplicação da DT proposta em uma base de dados de uma concessionária de energia, com a finalidade de diagnosticar a situação atual da planta, propondo assim acompanhamento preditivo das manutenções e intervenções a serem realizadas nos equipamentos.

## REFERÊNCIAS

- ALMEIDA, O. M.; AMORA, M. A. B.; BARBOSA, F. R.; ARAGAO, F. A. P.; NOTTINGHAM, O. C. E. S.; BARRETO, L. H. S. C.; VITOR, R. Monitoramento e Diagnóstico de um Transformador de Potência: Análise de Gases Dissolvidos no Óleo e Análise Térmica. *Anais do IV CITENEL - Congresso de Inovação Tecnológica em Energia Elétrica 2007*, Araxá, MG, 2007b.
- AMORA, M. A. B., ALMEIDA, O. M., BRAGA, A. P. S., BARBOSA, F. R., LIMA, S. S, LISBOA L. A. C. Extraction of Knowledge from Artificial Neural Networks to Application in Analysis of Transformers. The 15th International Conference on Intelligent System Applications to Power Systems, Curitiba-PR, 2009
- AMORA, M. A. B.; ALMEIDA, O. M.; BRAGA, A. P. S.; BARBOSA, PONTES, R. S. T. An improved DGA method based on rules extracted from a high-dimension input space. *Electronic Letters*, Volume 48, Issue 17, 16 August, p. 1048-1049, 2012.
- AMORA, M. A. B.; ALMEIDA, O. M.; BRAGA, A. P. S.; BARBOSA; R. S. T. PONTES. Rules extracted from a high-speed entrance space thought the addition of taxes to the obtain of a new method of analysis of incoming faults in transformers, Brazilian symposium on intelligent automation, 2013.
- ARANTES, J.G. Diagnóstico de falhas em transformadores de potência pela análise de gases dissolvidos em óleo isolante através de redes neurais. Dissertação (Mestrado em Ciências) – Universidade Federal de Itajubá, Itajubá, 2005.
- AZEVEDO, C. H. B. Metodologia para eficácia da detecção de descargas parciais por emissão acústica como técnica preditiva de manutenção em transformadores de potência imersos em óleo isolante. Dissertação de Mestrado, GPT/BC/UFG – Programa de Pós-Graduação em Engenharia Elétrica e de Computação / Universidade Federal de Goiás, Goiânia, GO, 2009.
- BARBOSA, F. R. Monitoramento on-line e diagnóstico inteligente da qualidade dielétrica do isolamento líquido de transformadores de potência, Dissertação de Mestrado, UFC/CT/PPGDEE – Programa de Pós-Graduação do Departamento de Engenharia Elétrica, Fortaleza, CE, 2008.
- BARBOSA, F.R., ALMEIDA, O. M., BRAGA, A. P.S., AMORA, M. A. B., CARTAXO, S. J. M. Application of an Artificial Neural Network in the Use of Physicochemical Properties as a Low Cost Proxy of Power Transformers DGA Data. *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 19, no.1, pp.239 – 246, 2012.
- BARBOSA, F.R. Diagnóstico de falhas incipientes a partir de propriedades físico-químicas do óleo isolante em transformadores de potência como método alternativo à análise de gases dissolvidos. Tese de Doutorado em Engenharia Elétrica, Universidade Federal do Ceará, Fortaleza, 2013.
- BASGALUPP, M.P. LEGAL-Tree: Um algoritmo genético multi-objetivo lexicográfico para indução de árvores de decisão. Tese de Doutorado, ICMC-USP, São Carlos, 2010.

BRACHMAN, R.; ANAND, T. The Process of Knowledge Discovery in Databases: A Human Centered Approach, in A KDDM, AAAI/MIT Press, 37-58, 1996.

BREIMAN, L.; FRIEDMAN, J.H.; OLSEN, R.A.; e STONE, C.J. Classification and Regression Trees. Wadsworth and Brooks, 1984.

CLAIBORNE, C. C.; CHERRY, D. B. A status update on the use of natural ester (vegetable oil) dielectric fluids in transformers. Anais do 74th Doble International Client Conference, Boston, MA, 2006.

CLARK, F. M. Insulating materials for design and engineering practice. s. ed. New York, John Wiley and Sons, v.1. p.131 -209, 1962.

COVER, T.M., "Geometrical and Statistical properties of systems of linear inequalities with applications in pattern recognition". IEEE Transactions on Electronic Computers EC-14: 326-334, 1965.

DÖRNENBURG, E.; GERBER, O. E. Analysis of Dissolved and Free Gases for Monitoring Performance of Oil-Filled Transformers. *The Brown Boveri Review*, v. 54, n. 213, pp. 104-111, 1967.

DUKARM, J. J. Transformer Oil Diagnosis Using Fuzzy Logic and Neural Networks. *Canadian Conference on Electrical and Computer Engineering*, v. 1, pp. 329-332, 1993.

DUVAL, M.; PABLO, A. de. Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases IEEE Elect. Insul. Mag., vol. 17, no. 2, pp. 31-41, 2001.

DUVAL, M. Review of Faults Detectable by Gas in Oil Analysis in Transformer IEEE Electrical Insulation Magazine, Vol. 18, N° 3, pp. 8-17, 2002.

DUVAL, M. "New techniques for dissolved gas in oil analysis" IEEE Elect. Insul. Mag., vol.19, no.2, pp. 6-15, 2003.

DUVAL, M.; DUKARM, J. "Improving the Reliability of Transformer Gas in Oil Diagnosis" IEEE Elect. Insul. Mag., vol.21, no.4, pp. 21-27, 2005.

DUVAL. "Dissolved Gas Analysis and the Duval Triangle" Fifth AVO New Zealand International Technical Conference, 2006.

ELDER, J.; PREGIBON, D. A Statistical Perspective on Knowledge Discovery in Databases, in *Advances in Knowledge Discovery and Data Mining*, Usama M. Fayyad, Gregory PiatetskyShapiro, Padhraic Smyth, and Ramasamy Uthurusamy (Eds.), AAAI Press, pp.83-113, 1996

FAYYAD, U. M.; e IRANI, K. B. The attribute selection problem in decision tree generation. in Proc. 10th Nat.Conf. Artificial Intelligence, Cambridge, MA, pp. 104–110, 1992.

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P. Data Mining to Knowledge Discovery: An Overview. In *Advances in Knowledge Discovery and Data Mining*, eds. U. Fayyad, G. PiatetskyShapiro, P. Smyth, and R. Uthurusamy, 1–30. Menlo Park, Calif.: AAAI Press, 1996.

FERREIRA, H.M. Uso de ferramentas de aprendizado de máquina para prospecção de perdas comerciais em distribuição de energia elétrica, Dissertação de Mestrado, FEEC/Unicamp, 2008.

FREUND, Y.; MASON, L. The alternating decision tree learning algorithm. In *Machine Learning: Proceedings of the Sixteenth International Conference*, 1999.

FRIEDENBERG, L. E. SANTANA, R. M. C. Propriedades de óleos isolantes de transformadores e a proteção do meio ambiente. In: SIMPÓSIO INTERNACIONAL DE QUALIDADE AMBIENTAL, 9., 2014, Porto Alegre. Anais... Porto Alegre-RS, 2014.

GOMEZ, N.A.G. Diagnóstico de falhas em transformadores isolados com óleo vegetal isolante a partir da análise de gases dissolvidos. Dissertação (Mestrado em Engenharia dos Materiais) – Setor de Tecnologia. Universidade Federal do Paraná, Curitiba, 2013.

GUARDADO, J. L.; NAREDO, J. L.; MORENO, P.; FUERTE, C. R. A comparative study of neural network efficiency in power transformers diagnosis using dissolved gas analysis. *IEEE Transactions on Power Delivery*, v. 16, n. 4, pp. 643 - 647, 2001.

HALSTEAD, W.D. A Thermodynamic Assessment of the Formation of Gaseous Hydrocarbons in Faulty Transformers. *Journal Institute of Petroleum*, v. 59, n. 9, pp. 239-41, 1973.

HAYKIN, S. *Redes Neurais: Princípios e Práticas*. Trad. Paulo Martins Engel, 2ª ed., Porto Alegre, RS, Bookman, 2001.

HELL, M.B. Sistemas de diagnóstico de faltas em transformadores de potência utilizando análise de gases dissolvidos e técnicas de inteligência computacional. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Católica de Minas Gerais, Belo Horizonte, 2002.

HOOSHMAND, R.A.; PARASTEGARI, M.; FORGHANI, Z. Adaptive neuro-fuzzy inference system approach for simultaneous diagnosis of the type and location of faults in power transformers. *Electrical Insulation Magazine*, IEEE Volume: 28. Issue: 5, pp 32-42, 2012.

HUANG, Y. C. Evolving neural nets for fault diagnosis of power transformers. *IEEE Transactions on Power Delivery*, v.18, n.3, pp. 843- 848, 2003.

HUANG, Y.C.; YANG, H.T.; HUANG, C.L. Developing a New Transformer Fault Diagnosis System through Evolutionary Fuzzy Logic. *IEEE Transactions on Power Delivery*, v. 12, n.2, pp. 761-767, 1997.

HUANG, Y. C, YANG, K., T; HUANG, K. Y. Abductive Network Model Based Diagnosis System for Power Transformer Incipient Fault Detection. *IEE Proceedings - Generation, Transmission and Distribution*. v. 149, n. 3, p. 326-330, May, 2002.

INTERNATIONAL ELECTROTECHNICAL COMMISSION – IEC 60599: Mineral Oil-Impregnated Electrical Equipment in Service - *Guide to the Interpretation of Dissolved and Free Gases Analysis*. Geneva, Switzerland, 2. ed., 1999.

INTERNATIONAL ELECTROTECHNICAL COMMISSION – IEC 60599: Mineral Oil-Impregnated Electrical Equipment in Service - *Guide to the Interpretation of Dissolved and Free Gases Analysis*. 2008.

IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers, IEEE Standard C57.104-2008, Sept. 2008.

IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers, IEEE Std C57.104-1991, 1991.

IZENMAN, A. J. Modern Multivariate Statistical Techniques. Springer, p45, 2008.

KAN, H.; MIYAMOTO, T. Proposals for an Improvement in Transformer Diagnosis Using Dissolved Gas Analysis (DGA). IEEE Electrical Insulation Magazine, vol. 11, no. 6, pp. 15-21, 1995.

KASS, G. V. An exploratory technique for investigating large quantities of categorical data. Appl. Statist., vol. 29, no. 2, pp. 119–127, 1980.

KOHAVI R., and PROVOST F., Glossary of Terms, Machine Learning 30(2/3): 271-274,1998.

LANDWEHR, N., HALL, M., & FRANK, E. Logistic model trees. Machine Learning, 59(1-2):161-205, 2005.

LI, X.; e WU, H. DGA interpretation scheme derived from case study. IEEE Trans. Power Delivery, vol. 26, no. 2, pp. 1292 – 1293, April 2011.

LIMA, S. E. U.; ALMEIDA, O. M.; REIS, L. L. N. Diagnóstico de Falhas Incipientes em Transformadores de Potência Imersos em Óleo Utilizando Sistemas Nebulosos. XV Congresso Brasileiro de Automática (CBA 2004), Gramado, RS, 2004.

LIMA, S. E. U. Diagnóstico Inteligente de Falhas Incipientes em Transformadores de Potência Utilizando a Análise dos Gases Dissolvidos em Óleo. *Dissertação de Mestrado*, UFC/CT/DEE - Departamento de Energia Elétrica, Fortaleza, CE, 2005.

LIMA, S. E. U.; BARBOSA, F. R.; ALMEIDA, O. M. Diagnóstico de Falhas Incipientes em Transformadores de Potência. *Simpósio Brasileiro de Sistemas Elétricos (SBSE 2006)*, Campina Grande, PB, 2006.

LIMA, S. L. de. Diagnóstico de falhas incipientes de transformadores de potência utilizando a teoria da expansão. *Dissertação de Mestrado*, UFMA, São Luís, MA, 2008.

LIPSHTEIN, R. A.; SHAKHNOVICH, M. I. Transformer oil. 2. ed. Jerusalem, Israel, p.275, 1970.

LOH, T.; e SHIH, T. Split selection methods for classification trees. Statistica Sinica, vol. 7, pp. 815–840, 1997.

MILASCH, M. Manutenção de Transformadores em Líquido Isolante. *Edgar Blücher*, São Paulo, SP, 1984.

MIRANDA, V.; CASTRO, A. R. G. Improving the IEC Table for Transformer Failure Diagnosis with Knowledge Extraction from Neural Networks. *IEEE Transactions on Power Delivery*, v. 20, n. 4, pp. 2509-2516, 2005.

- MORAIS, D. R. Ferramenta Inteligente para Detecção de Falhas Incipientes em Transformadores Baseada na Análise de Gases Dissolvidos no Óleo Isolante. *Dissertação de Mestrado*. UFSC, Florianópolis, SC, 2004.
- MOURA, C.R.S.; CARRETEIRO, R.P. Lubrificantes e lubrificação. Rio de Janeiro, Livros Técnicos e Científicos Editora S.A , 1975.
- MYERS, S. D.; KELLY, J. J.; PARRISH, R. H. A guide to transformer maintenance. Ohio: Akron, p830, 1981.
- NARESH, R., SHARMA, V. and VASHISTH, M. "An integrated neural fuzzy approach for fault diagnosis of transformers," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 2017-2024, 2008.
- NIBLETT, T.; BRATKO, I. Learning decision rules in noisy domains. in *Expert Systems*. Cambridge, MA: Cambridge Univ. Press, 1986.
- NYNÄS NAPHTHENICS AB. Transformer oil handbook. 1 ed., Sweden: Linderoths in Vingåker, 223 p, 2004.
- NOGUEIRA, C.A.A. Avaliação de sistemas de medição da concentração de hidrogênio dissolvido em óleo isolante. *Dissertação (Mestrado em Metrologia) – Universidade Federal de Santa Catarina, Florianópolis, 2004.*
- PATEL, N. K.; KHUBCHANDANI, R. K. ANN Based Power Transformer Fault Diagnosis. *IE Journal*, v. 85, n. 1, p. 60-63, 2004.
- QUINLAN, J.R. Induction of Decision Trees. *Mach. Learn.* 1, pp.81-106, 1986.
- QUINLAN, J. Decision trees and multivalued attributes. *Machine Intelligence*, 11:305-318, 1988.
- QUINLAN, J.R. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, 1993.
- QUINLAN, J. Learning decision tree classifiers. *ACM Computing Surveys (CSUR)*, 28(1):71–72, 1996.
- RANDY, D. S. et al., Power Transformer Fault Diagnosis, IEEE PES WM. Panel Session, New York, Feb 6, 1977.
- RAKATOMALA, R. SIPINA. Programa computacional. Disponível em <<http://eric.univ-lyon2.fr/~ricco/sipina.html>>. Acesso em Outubro de 2017.
- ROGERS, R.R. IEEE and IEC Codes to Interpret Incipient Faults in Transformers Using Gas in Oil Analysis. *IEEE Transactions Electrical Insulation*, v. 13, n. 5, p. 349-354, 1978.
- ROKACH, L., and MAIMON, O. Top-down induction of decision trees classifiers - a survey. *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 35, no 4, pp. 476 – 487, Nov. 2005.
- ROKACH, L., and MAIMON, O. Data mining with decision trees: Theory and Applications. *Series in Machine Perception and Artificial Intelligence - Vol. 69*. World Scientific Publishing, 2008.

ROKACH, L., and MAIMON, O. Data mining with decision trees: Theory and Applications. Series in Machine Perception and Artificial Intelligence - Vol. 81. World Scientific Publishing, 2014.

RUSSELL, S.; NORVIG, P. Inteligência Artificial. Editora Campus, 2004.

SENNA, S.L. Computação evolucionária aplicada ao diagnóstico de falhas incipientes em transformadores de potência utilizando dados de cromatografia. Dissertação (Mestrado em Engenharia Elétrica). Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

SFERRA, H.; CORRÊA, Â. M. J. Conceitos e aplicações de data mining. Revista de ciência & tecnologia, 11:19–34, 2003.

SHERIF, S.M.G and IBRAHIM, B.M.T. A new approach of DGA interpretation technique for transformer fault diagnosis, *IEEE Electrical Power and Energy Systems* 81(2016), pp. 265-274, 2016.

SHI, H. (2007). Best-first decision tree learning. Master's thesis, University of Waikato, Hamilton, NZ, 2007.

SIADA, A.A. and HMOOD, S. A new fuzzy logic approach to identify power transformer criticality using dissolved gas-in-oil analysis, *IEEE Electrical Power and Energy Systems* 67, pp. 401-408, 2015.

SILVA, I. N.; SOUZA, A. N.; SOUZA, M. G. Z. N. Avaliação do Nível de Contaminação do Óleo de Transformadores Através de Redes Neurais Artificiais. In: XV SNPTEE, GSE 13. Anais. Paraná, Brasil. Outubro, 1999.

SINGH, S.; BANDYOPADHYAY, M.N. Dissolved gas analysis technique for incipient fault diagnosis in power transformers: A bibliographic survey. *IEEE Electrical Insulation Magazine*, v. 26, n. 6, pp. 41-46, 2010.

SOUSA, R.; AMORA, M.; BARBOSA, F. R.; PONTES, R. T.; BRAGA, A. P. S.; ALMEIDA, O. M. "Análise de Métodos para o Diagnóstico de Falhas Incipientes em Transformadores com Base na Concentração de Gases no Óleo Isolante". IV Simpósio Brasileiro de Sistemas Elétricos - SBSE 2012, Goiás, 2012.

STOCCO, M. Avaliação do potencial de aplicação de óleos vegetais como fluidos isolantes em transformadores de distribuição da rede elétrica. 2009. 129 f. Dissertação (Mestrado) - Curso de Engenharia e Ciência dos Materiais, Universidade Federal do Paraná, Curitiba. 2009.

THEODORIDIS, S.; KOUTROUMBAS, K. Pattern Recognition. 3. ed. San Diego: Academic Press, 2009.

WANG, Z.; LIU, Y.; GRIFFIN, P.J. A combined ANN and expert system tool for transformer fault diagnosis. *IEEE Transactions on Power Delivery*, v. 13, n. 4, pp. 1224-1229, 1998.

WANG, Z. Artificial Intelligence Applications in the Diagnosis of Power Transformer Incipient Faults. *PhD Thesis*, Virginia Polytechnic Institute and State University, Blacksburg, USA, 105 p, 2000.

WILSON, G. W.; McLean, A.; Dessulfurization of Iron and Steel and Sulfide Shape Control, The Iron & Steel Society of AIME, Warrendale, pp.1-41, USA, 1980.

XU, W., WANG, D., ZHOU, E. Z.; CHEN, H. Fault Diagnosis of Power Transformers: Application of Fuzzy Set Theory, Expert Systems and Artificial Neural Networks. IEE Proc. – Sci. Meas. Technol., v. 144, n. 1, Jan. 1997.

ZHANG, Y. An Artificial Neural Network Approach to Transformer Fault Diagnosis, *Thesis*, Virginia Tech, Blacksburg, USA, 1996.

ZIRBES, R.; ROLIM, J. G.; ZÜRN, H. H. Metodologias para Avaliação e Diagnóstico do Estado de Isolamentos de Papel Impregnado com Óleo Mineral. *SBA Controle & Automação*, vol. 16, no. 3, pp. 318-331, 2005.