

**UNIVERSIDADE FEDERAL DE PERNAMBUCO**  
**CENTRO DE TECNOLOGIA E GEOCIÊNCIAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**Ferramenta Inteligente de Classificação do Diagnóstico  
de Falhas Incipientes em Transformadores de Potência  
Imersos em Óleo**

por

**AGNALDO MAGNUM DE SOUZA SILVA**

Dissertação (Tese) submetida ao Programa de Pós-Graduação em Engenharia Elétrica da  
Universidade Federal de Pernambuco como parte dos requisitos para a obtenção do grau de  
Mestre em Engenharia Elétrica.

**ORIENTADORA: MILDE MARIA DA SILVA LIRA, D. Sc.**  
**CO-ORIENTADOR: RONALDO RIBEIRO BARBOSA DE AQUINO, D. Sc.**

Recife, fevereiro de 2012.

©Agnaldo Magnum de Souza Silva, 2012

Resumo da Dissertação apresentada à UFPE como parte dos requisitos necessários  
para a obtenção do grau de Mestre em Engenharia Elétrica.

# **Ferramenta Inteligente de Classificação do Diagnóstico de Falhas Incipientes em Transformadores de Potência Imersos em Óleo**

Agnaldo Magnum de Souza Silva

Fevereiro/2012

Orientadora: Milde Maria da Silva Lira, D. Sc.

Co-Orientador: Ronaldo Ribeiro Barbosa de Aquino, D. Sc.

Área de Concentração: Processamento de Energia

Palavras-chave: Falhas Incipientes, Transformador de Potência, Lógica Fuzzy, Óleo Isolante.

Número de Páginas: 78

A análise dos gases dissolvidos no óleo isolante (DGA) é uma das técnicas mais utilizadas para detectar faltas incipientes em transformadores de potência. A identificação das faltas ocorrendo no transformador nem sempre é uma tarefa simples devido à variação dos gases presentes no óleo ao longo da operação do equipamento. O trabalho tem por objetivo apresentar um sistema inteligente de manutenção preditiva que deve detectar automaticamente, a partir da análise cromatográfica do óleo isolante, qual o tipo de falha incipiente mais provável de está ocorrendo no transformador de potência. Essa análise é baseada em normas internacionais que são largamente utilizadas pelas empresas do setor elétrico e a modelagem do sistema foi realizada utilizando técnicas de inteligência artificial, como Redes Neurais e Lógica Fuzzy. Os dados utilizados para o desenvolvimento da pesquisa foram fornecidos pela Companhia Energética de Pernambuco (CELPE) e possuem análises coletadas desde 1985 até 2010 de transformadores distribuídos em subestações existentes em todo o estado. Os resultados obtidos mostraram-se superiores quando comparados com os resultados de pesquisas que abordaram o mesmo tipo de problema.

Abstract of Dissertation presented to UFPE as a partial fulfillment of the requirements for the degree of Master in Electrical Engineering.

## **Intelligent Tool rating Incipient Fault Diagnosis in Power Transformers Immersed in Oil**

Agnaldo Magnum de Souza Silva

February/2012

Supervisors: Milde Maria da Silva Lira, D. Sc.;Ronaldo Ribeiro Barbosa de Aquino, D. Sc.

Area of Concentration: Energy Processing

Keywords: Incipient Fault, Power Transformers, Fuzzy Logic, Insulating Oil.

Number of Pages: 78

The dissolved gas analysis (DGA) is one of the most widely used techniques to detect incipient faults in power transformers. The identification of faults occurring in the transformer is not always a simple task due to the variation of the gases present in the oil over the operation of the equipment. The paper is intended to provide an intelligent system for predictive maintenance that should automatically detect, starting from gas chromatographic analysis of insulating oil, which type of incipient fault is more likely to occur in the power transformer. This analysis is based on international standards that are widely used by electric power companies and system modeling was performed using artificial intelligence techniques, such as Neural Networks and Fuzzy Logic. The data used for the development of the research were provided by Companhia Energética de Pernambuco (CELPE) and analyses have collected from 1985 to 2010 of transformers in existing substations distributed throughout the state. The results proved to be superior when compared with the results of research that addressed the same issue.

## Sumário

<b>Lista de Tabelas</b> .....	<b>vi</b>
<b>Lista de Figuras</b> .....	<b>vii</b>
<b>Lista de Equações</b> .....	<b>viii</b>
<b>Siglas</b> .....	<b>x</b>
<b>1.Introdução</b> .....	<b>1</b>
1.1. Motivação .....	1
1.2. Técnicas adotadas para diagnóstico .....	2
1.3. Divisão da dissertação.....	3
<b>2.Redes Neurais Artificiais</b> .....	<b>5</b>
2.1. Funcionamento do neurônio .....	6
2.2. Funções de ativação .....	7
2.3. Treinamento das redes neurais.....	8
2.3.1. Gradiente descendente (GD) .....	9
2.3.2. Algoritmo Resilient Propagation (Rprop) .....	11
2.3.3. Algoritmo Levenberg-Maquardt .....	12
<b>3.Teoria dos Conjuntos Fuzzy</b> .....	<b>14</b>
3.1. Propriedades e operações .....	17
3.2. Variáveis linguísticas .....	18
3.3. Funções de pertinência.....	19
3.4. Sistema de Inferência Mamdani .....	22
3.5. Sistema de Inferência Takagi-Sugeno .....	23
<b>4.Óleo Isolante</b> .....	<b>25</b>
4.1. Análise dos gases dissolvidos no óleo isolante.....	26
4.1.1. Método IEC .....	29
4.1.2. Método de ROGERS .....	33
<b>5.Implementação Redes Neurais</b> .....	<b>36</b>
5.1. Conjunto de dados .....	36
5.1.1. Pré-processamento dos dados.....	36
5.1.2. Seleção de Atributos.....	37
5.1.3. Elaboração do conjunto de dados para utilização na Rede Neural....	38
5.2. Arquitetura da rede neural .....	42
5.2.1. Rede Multilayer Perceptron – MLP .....	42
5.2.2. Determinação do número de nodos da camada intermediária.....	44

<b>6.Implementação Lógica Fuzzy.....</b>	<b>49</b>
6.1. Desenvolvimento do SIF .....	50
6.1.1. Método IEC .....	50
6.1.2. Método ROGERS .....	54
<b>7.Resultados .....</b>	<b>57</b>
7.1. Lógica Fuzzy <i>versus</i> Rede Neural .....	57
7.2. Exemplos de resultados .....	60
<b>8.Conclusão .....</b>	<b>63</b>
<b>Revisão Bibliográfica .....</b>	<b>65</b>

## Lista de Tabelas

Tabela 1 - Propriedades algébricas dos conjuntos.....	17
Tabela 2 – Operações e Relações com Conjuntos Fuzzy .....	18
Tabela 3 – Codificação das relações dos gases pelo método da IEC .....	31
Tabela 4 – Diagnósticos sugeridos pelo método da IEC .....	31
Tabela 5 – Avaliações da relação CO <sub>2</sub> /CO.....	33
Tabela 6 – Construção método ROGERS .....	34
Tabela 7 – Diagnósticos sugeridos .....	35
Tabela 8 - Codificação dos dias da semana.....	38
Tabela 9 – Quantidade de diagnóstico por classe para o método IEC. ....	39
Tabela 10 - Quantidade de diagnóstico por classe para o método ROGERS.....	40
Tabela 11 – Código do vetor de saída da rede neural.....	44
Tabela 12 – Erro médio quadrado das redes.....	45
Tabela 13 – Erro percentual de classificação no conjunto de teste .....	46
Tabela 14 – Percentual de acerto na classificação.....	46
Tabela 15 - Codificação IEC .....	51
Tabela 16 – Percentual de acerto na classificação por diagnóstico.....	57
Tabela 17 – Resultados IEC .....	58
Tabela 18 – Resultados ROGERS .....	59
Tabela 19 – Acompanhamento de diagnóstico do Transformador.....	60
Tabela 20 – Exemplo de aplicação da metodologia .....	61

## Lista de Figuras

Figura 1 – Neurônio artificial .....	7
Figura 2 – Funções de ativação. ....	8
Figura 3 - Funções de ativação mais utilizadas. ....	8
Figura 4 - Funções de pertinência para a variável estatura.....	18
Figura 5 - Diferentes formatos das funções de pertinência - triangular (a), trapezoidal (b), Gaussiana (c), conjunto unitário (d) .....	20
Figura 6 – Funções de pertinência .....	21
Figura 7 – Relação entre Temperatura e Geração de Gases no Óleo Isolante[19].....	26
Figura 8– Representação gráfica das razões dos gases [3].....	32
Figura 9 – Esquema de uma rede MLP com duas camadas intermediárias .....	43
Figura 10- Topologia da rede MLP selecionada .....	48
Figura 11 – Funções membro para a relação R1 .....	51
Figura 12 – Representação da metodologia para $R_1=2,9$ . ....	52
Figura13 - Funções membro para a relação R2.....	53
Figura 14 - Funções membro para a relação R3 .....	53
Figura 15– Funções membro para R1 ROGERS.....	54
Figura 16 – Funções membro para R2 ROGERS.....	55
Figura 17 – Funções membro para R3 ROGERS.....	56
Figura 18 – Funções membro para R4 ROGERS.....	56

## Lista de Equações

Equação ( 1 ).....	9
Equação( 2 ).....	9
Equação ( 3 ).....	10
Equação ( 4 ).....	10
Equação ( 5 ).....	10
Equação ( 6 ).....	10
Equação ( 7 ).....	10
Equação ( 8 ).....	10
Equação ( 9 ).....	11
Equação ( 10 ).....	12
Equação ( 11 ).....	12
Equação ( 12 ).....	12
Equação ( 13 ).....	13
Equação ( 14 ).....	13
Equação ( 15 ).....	13
Equação ( 16 ).....	13
Equação ( 17 ).....	13
Equação ( 18 ).....	13
Equação ( 19 ).....	14
Equação ( 20 ).....	15
Equação ( 21 ).....	15
Equação ( 22 ).....	15
Equação ( 23 ).....	15
Equação ( 24 ).....	16
Equação ( 25 ).....	16
Equação ( 26 ).....	16
Equação ( 27 ).....	17
Equação ( 28 ).....	19
Equação ( 29 ).....	20
Equação ( 30 ).....	20
Equação ( 31 ).....	20

Equação ( 32 ).....	20
Equação ( 33 ).....	22
Equação ( 34 ).....	22
Equação ( 35 ).....	23
Equação ( 36 ).....	23
Equação ( 37 ).....	24
Equação ( 38 ).....	24
Equação ( 39 ).....	37
Equação ( 40 ).....	38
Equação ( 41 ).....	39
Equação ( 42 ).....	52
Equação ( 43 ).....	52
Equação ( 44 ).....	52
Equação ( 45 ).....	52
Equação ( 46 ).....	52
Equação ( 47 ).....	52
Equação ( 48 ).....	52
Equação ( 49 ).....	52
Equação ( 50 ).....	52
Equação ( 51 ).....	54
Equação ( 52 ).....	54
Equação ( 53 ).....	54
Equação ( 54 ).....	55
Equação ( 55 ).....	55
Equação ( 56 ).....	55
Equação ( 57 ).....	55
Equação ( 58 ).....	55
Equação ( 59 ).....	55
Equação ( 60 ).....	55
Equação ( 61 ).....	55

## Siglas

IA	Inteligência Artificial
LF	Lógica Fuzzy
RNA	Rede Neural Artificial
IEC	International Eletrotechnical Comission
CELPE	Companhia Elétrica de Pernambuco
DGA	Dissolves Gas Analisis
GD	Gradiente Descedente
Rprop	Resilient Propagation
<i>GDM</i>	Gradiente Descendente Com Momento
MLP	Rede Multilayer Perceptron
MSE	Mean Square Error
SIF	Sistema de Inferência Fuzzy

# 1. Introdução

## 1.1. Motivação

Considerando o modelo atual do setor elétrico nacional e a necessidade de redução dos seus custos operacionais no sentido de viabilizar a própria sobrevivência da empresa num mercado cada vez mais competitivo e globalizado, estamos assistindo a um avanço na tecnologia utilizada e nos processos de automatização das instalações, com destaques, por exemplo, para a monitoração e diagnóstico de estado dos equipamentos de forma preditiva (on-line), ou seja, sem indisponibilizar o equipamento, como os transformadores de potência, considerados como os principais equipamentos dentro da distribuição de energia.

Embora este evento tenha chegado de forma definitiva, observamos que a aplicabilidade do mesmo de forma generalizada ainda está muito distante, principalmente devido ao fator econômico, pois um sistema completo de monitoramento e diagnóstico de estado associado a um transformador de potência pode representar um custo adicional na ordem de 20 % (vinte por cento) do preço do equipamento[1] acrescido ainda da falta de dados históricos nesses sistemas de diagnósticos para a emissão confiável de pareceres conclusivos sobre o real estado operacional do transformador.

Assim sendo é de suma importância que cada empresa, com base na análise de desempenho operacional dos seus transformadores, defina as suas prioridades, não só em função da importância relativa do equipamento para o sistema, mas principalmente, optando na monitoração apenas daqueles parâmetros que possam diagnosticar incipientemente um tipo de problema no equipamento com maior probabilidade para algumas famílias de equipamentos e que, dentro de um curto espaço de tempo, possa o defeito evoluir para uma falha.

Os transformadores são equipamentos de extrema importância numa subestação elétrica, pois além de possuírem um custo bastante elevado a sua indisponibilidade pode causar o desligamento de toda uma rede de distribuição, provocando um desgaste para imagem da empresa e podendo gerar o pagamento elevadas multas pela não prestação do serviço. Analisando esse aspecto as técnicas que monitoram o desempenho desses equipamentos sem retirá-los de operação devem ser largamente pesquisadas com o intuito de melhorar os diagnósticos fornecidos e evitar que defeitos mais graves possam inutilizar o equipamento.

## 1.2. Técnicas adotadas para diagnóstico

Através dos gases dissolvidos no óleo isolante do transformador é possível avaliar estado do seu isolamento com a indicação de possíveis estados de falhas que possam estar ocorrendo. Esses processos de análise de gases dissolvidos (DGA) ou análise cromatográfica do óleo isolante são realizados a partir da amostra do óleo isolante de um transformador em operação, quantificando a concentração de certos gases gerados no equipamento e dissolvidos no óleo isolante, indicando a presença de falhas incipientes associadas à parte ativa do transformador.

Os métodos de análise cromatográfica do óleo isolante são largamente utilizados pelas companhias de distribuição elétrica

Existem diversas ferramentas que apontam a interpretação da análise cromatográfica do óleo isolante, dentre as quais podem ser citadas:

- Método Pugh ou pelo gás chave;
- Método pela IEC ([2] e [3])
- Método Rogers ([4] e [5])
- Método Duval;
- Método Dornenburg;

Esses métodos ou critérios são caracterizados por experiências, muitas vezes empíricas e associa a formação de alguns gases chaves ao tipo de falha incipiente mais provável de está ocorrendo internamente ao transformador.

Durante o desenvolvimento deste trabalho muitos destes métodos foram pesquisados, porém somente os mais representativos e com disponibilidade de dados comparativos foram utilizados para a elaboração da ferramenta desenvolvida.

Em pesquisas recentes, verifica-se a possibilidade de desenvolvimento de um modelo computacional utilizando técnicas de Inteligência Artificial (IA), tais como, Lógica Fuzzy (LF) e Redes Neurais Artificiais (RNA) para realizar o diagnóstico preditivo de falhas em transformadores de potência.

A RNA é treinada satisfatoriamente e indica um diagnóstico que é similar aos diagnósticos das amostras já definidos pelos métodos IEC e ROGERS. Depois, pode-se utilizar a mesma rede neural para definir os prováveis diagnósticos para as amostras que inicialmente não tinham diagnósticos tabelados pelo método IEC e/ou ROGERS. A Lógica

Fuzzy além de fornecer um possível defeito para os casos tabelados nos métodos IEC e ROGERS, também indica uma provável falha não tabelada pelas normas.

A eficiência no resultado obtido pela lógica fuzzy para as amostras com diagnósticos não tabelados foi feita baseando-se em casos conhecidos comparando-se o resultado obtido com o informado nos laudos fornecidos pela CELPE.

Assim, tomou-se como ponto de partida um sistema desenvolvido em RNA e posteriormente investigado com outra técnica de IA, como Lógica Fuzzy.

Sendo assim, o sistema computacional desenvolvido é capaz de repetir os resultados da norma IEC e ROGERS para os diagnósticos definidos. Desta forma, a empresa teria condições de atuar preventivamente nos seus equipamentos aumentando a disponibilidade e a confiabilidade além de reduzir os custos de manutenção, pois atuando em falhas incipientes poderia fazer manutenções programadas que teriam custo de reparo menor, já que o equipamento não estaria totalmente danificado. Ou seja, o defeito não evoluiu para uma falha com saída forçada do equipamento.

Atualmente a CELPE já realiza o ensaio de análises de gases dissolvidos – DGA, sendo estes os dados que serão utilizados no sistema aqui proposto.

Dentro dessa análise, o critério utilizado pela CELPE com maior êxito é o critério de diagnóstico pela IEC, seguido depois pelo critério de ROGERS. Com base na bibliografia e casos de outras empresas, constatou-se que cerca de 70% (setenta por cento) do total de amostras de análises cromatográficas estão com diagnósticos definidos entre os 10 (dez) tipos indicados pela IEC (tabelados), e 30% (trinta por cento) não apresentam diagnóstico pela IEC (não tabelados).

Diante do exposto está justificada a necessidade e importância do projeto bem como a sua originalidade baseada na utilização de técnicas de Inteligência artificial, Redes Neurais, Lógica Fuzzy e Algoritmos genéticos além de regras heurísticas.

### **1.3.Divisão da dissertação**

Este trabalho foi dividido em capítulos para um melhor entendimento e organização das ideias.

O capítulo 2 traz as teorias que envolvem a manipulação das redes neurais, como funcionam os neurônios artificiais e os principais algoritmos de treinamento utilizados.

O capítulo 3 explica o funcionamento da lógica fuzzy e mostra as suas principais propriedades, incluindo as operações possíveis de realizar para utilização dessa ferramenta.

A importância e a metodologia de análise do gás presente no óleo isolante é mostrada no capítulo 4, onde também são descritos os métodos de análise adotados ao longo do desenvolvimento da pesquisa.

Os métodos adotados para a criação das ferramentas propostas são explicados nos capítulos 0 e 6, que traz a metodologia implantada para a criação da rede neural e lógica fuzzy, respectivamente.

O capítulo 7 mostra uma comparação entre as ferramentas elaboradas, além de exemplificar alguns resultados reais que puderam ser extraídos da criação da ferramenta.

O capítulo 8 traz a conclusão do trabalho dissertando sobre os principais objetivos atingidos ao longo dos capítulos do trabalho.

## 2. Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) são sistemas paralelos distribuídos compostos por unidades de processamentos simples que computam determinadas funções matemáticas. Estas unidades estão dispostas em uma ou mais camadas conectadas entre si, geralmente, unidirecionais. Na maioria dos modelos estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede. O funcionamento destas redes é inspirado em uma estrutura física inspirada no cérebro humano[6].

Redes neurais artificiais vêm sendo desenvolvidas como generalização de modelos matemáticos do conhecimento humano, baseado nas suposições presentes em[7], listadas abaixo.

- O processamento da informação ocorre através de muitos elementos simples, chamados neurônios;
- Os sinais são transmitidos entre os neurônios através de um elo de conexão;
- Cada elo de conexão tem um peso associado, o qual, em redes neurais típicas, é múltiplo do sinal transmitido;
- Cada neurônio aplica uma função de ativação (normalmente não linear) em sua rede de entrada (soma dos pesos multiplicados pelos sinais de entrada) para determinar seu sinal de saída.

As redes neurais são caracterizadas por:

- Seu padrão de conexão entre os neurônios (chamado arquitetura),
- Seu método de determinação dos pesos nas conexões (chamado treinamento ou aprendizado), e
- Sua função de ativação[7].

Cada um destes elementos será explicado em mais detalhes ao longo do trabalho. A rede neural consiste num grande número de elementos de processamento simples chamados neurônios, unidades, células ou nós. Cada neurônio é conectado a outro neurônio por meio de elos de comunicação direta, cada qual com um peso associado. Os pesos representam a informação que será usada pela rede para resolver o problema. As redes neurais podem ser aplicadas a uma grande variedade de problemas, tal como

armazenagem e recuperação de dados, classificação de padrões, mapeamento de padrões de entrada para padrões de saída, agrupamento de padrões similares, ou para encontrar soluções em problemas de otimização com restrições [7].

Os neurônios possuem estados de ativação, que é função do seu sinal de entrada, ao modificar o seu estado ele envia sua ativação como um sinal para os vários outros neurônios. Apesar de só poder enviar um sinal por vez esse sinal é enviado a todos os outros neurônios. De forma que uma rede neural é composta por diversas unidades de processamento, onde cada unidade possui um funcionamento bastante simples. Essas unidades geralmente estão conectadas por canais de comunicação que estão associados a um determinado peso.

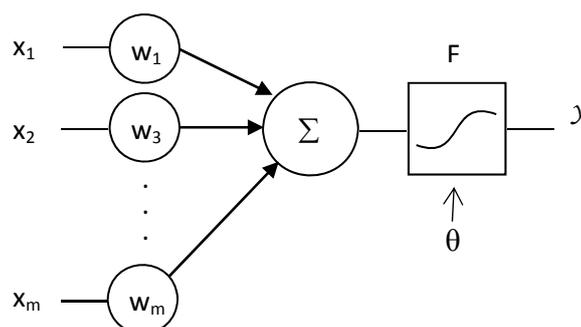
Cada unidade faz operação somente sobre os seus dados locais, os quais são entradas recebidas através de suas conexões. O comportamento inteligente atribuído as Redes Neurais Artificiais é originário das interações entre cada unidade de processamento. Tipicamente a arquitetura de uma rede neural está dividida em camadas onde as unidades de uma camada poderão estar conectadas às unidades da camada posterior.

As Redes podem ser classificadas pela arquitetura em que se apresentam, topologias, características dos seus nós, regras de treinamento e tipos de modelo.

## **2.1. Funcionamento do neurônio**

Como já dito o neurônio artificial é um dispositivo bastante simples, ele deve ser capaz de calcular o somatório de  $N$  entradas ponderadas, onde o resultado deverá ser apresentado a uma função não linear.

A Figura 1 exibe a representação de um neurônio artificial onde os valores  $x_1, x_2, \dots, x_n$  representam as entradas do neurônio as quais serão ponderadas pelos pesos  $w_1, w_2, \dots, w_n$  e a saída  $y$  que será dada pelo somatório das entradas ponderadas e passadas por uma função não linear, chamada função de ativação.



**Figura 1** – Neurônio artificial

Um neurônio de uma rede qualquer enviará o impulso para um próximo neurônio caso os sinais recebidos sejam suficientes para estimular a sua função de ativação acima do seu limiar, chamado Bias, caso as entradas não atingirem este nível o sinal decairá e não gerará qualquer ação.

Cada entrada do neurônio é multiplicada pelo peso associado a sua conexão ( $X_i W_i$ ). Dependendo da função de ativação, se o peso é positivo ele ajuda na excitação do nó de saída, já os pesos negativos tendem a inibir a ativação da função de saída.

O valor de  $\theta$  tem por finalidade aumentar o número de grau de liberdade do modelo, permitindo um aumento da capacidade da rede neural de se ajustar o conhecimento que lhe é fornecido, esse valor afeta a ativação do nó de saída  $y$ .

## 2.2. Funções de ativação

As funções de ativação de uma rede neural executam uma operação matemática no sinal de saída, podem ser utilizadas funções bastante sofisticadas a depender da complexidade do problema a ser resolvido pela rede, mais comumente são utilizadas as seguintes funções mostradas na Figura 2.

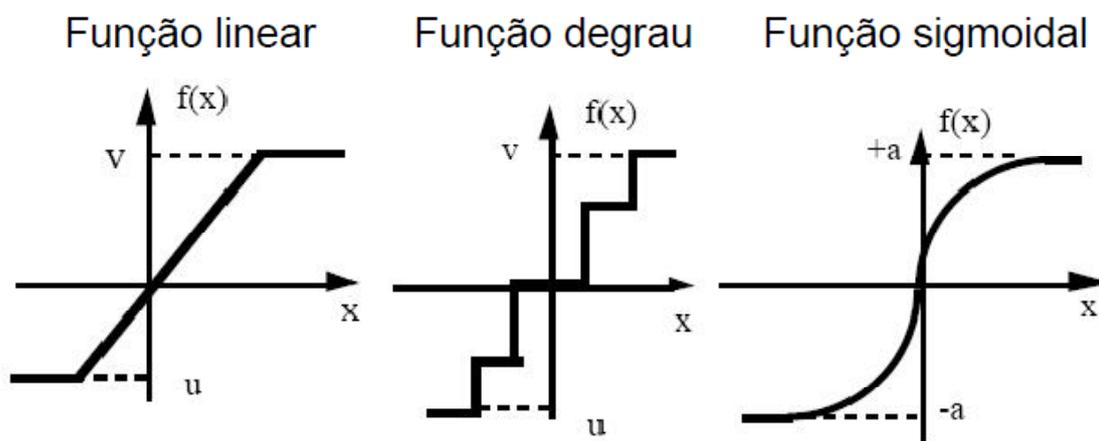


Figura 2 – Funções de ativação.

Entre as funções de ativação mais utilizadas, destaca-se a função sigmoidal que é a função de ativação mais comum de ser utilizada em redes neurais, entre as mais utilizadas dessa função estão a sigmoide logística e a tangente hiperbólica[8], mostradas na Figura 3.

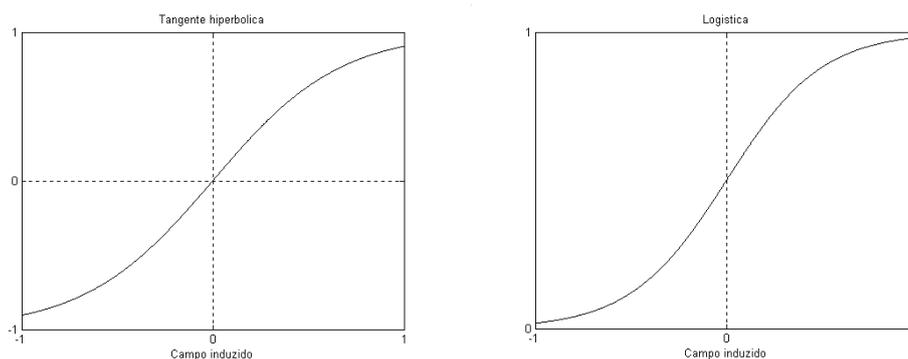


Figura 3 -Funções de ativação mais utilizadas.

Normalmente a função de ativação é definida pelo idealizador da rede e os parâmetros ajustados, de acordo com uma regra de aprendizagem, com as relações de entrada/saída para um cada caso específico.

### 2.3. Treinamento das redes neurais

O algoritmo de aprendizado supervisionado com retropropagação de erro (*errorbackpropagation*) ([8]e[9]) é utilizado para realizar o treinamento das redes neurais MLP. O algoritmo de retropropagação de erro possui duas fases distintas:

- Para frente ou propagação (em inglês *forward*). As características de uma amostra são introduzidas na camada de entrada, propaga-se para frente (neurônio por neurônio) através da rede e emerge na camada de saída da rede;
- Para trás ou retropropagação (em inglês *backward*). Os erros que se originam nos neurônios da camada de saída se propagam para trás (camada por camada) por meio da rede, atualizando os seus pesos.

Existem vários métodos de atualização dos pesos destacamos: o método do gradiente descendente (GD), *Resilient Propagation* (Rprop) e o *Levenberg-Maquart* (LM). Esta pesquisa utilizou os dois últimos métodos.

### 2.3.1. Gradiente descendente (GD)

O algoritmo do gradiente descendente é baseado no algoritmo de retropropagação de erro (criado a partir de uma generalização da regra delta) sendo composto por duas fases: a fase *forward*, na qual um padrão de dados é apresentado à rede e esta processa os dados, produzindo respostas, e a fase *backward*, que utiliza o erro obtido entre o resultado da execução *forward* e o resultado desejado (que é conhecido, de onde vem o termo *treinamento supervisionado*), para determinar os ajustes a serem feitos nos pesos (conexões sinápticas) dos neurônios da rede.

O algoritmo de retropropagação é apresentado a seguir.

O sinal de erro na saída do neurônio  $j$ , na iteração  $n$ , é definido pela Equação ( 1 ).

$$e_k(n) = d_k(n) - y_k(n) \quad (1)$$

onde,  $d_k(n)$  é a saída desejada e  $y_k(n)$  é a saída apresentada na saída do neurônio.

A função de custo  $\varepsilon$  é a energia total do erro, obtida somando o quadrado dos erros de todos os neurônios da camada de saída, Equação ( 2 ).

$$\varepsilon(n) = \frac{1}{2} \sum_{k=1}^m e_k^2(n) \quad (2)$$

onde  $m$  é o número de neurônios da camada de saída da rede e  $e_k$  é o erro do neurônio  $k$ , na iteração  $n$ .

O cálculo dos deltas para a correção dos pesos depende da posição da camada. Considere o neurônio da camada de saída  $k$ , sendo estimulado por um conjunto de ativações produzido por uma camada de neurônios à sua esquerda. A saída  $v(n)$  do neurônio  $k$  é, portanto:

$$v_k(n) = \sum_{j=0}^{m_2} w_{kj}(n) y_j(n) \quad (3)$$

Representando a Função de ativação  $y_k(n)$  que aparece na saída do neurônio  $k$  na iteração  $n$  é:

$$y_k(n) = \varphi_k(v_k(n)) \quad (4)$$

O algoritmo de retropropagação aplica uma correção  $\Delta w_{kj}$  ao peso sináptico  $w_{kj}$ , que é proporcional à derivada parcial  $\partial \varepsilon(n) / \partial w_{kj}(n)$ . De acordo com a regra da cadeia do cálculo, se expressa este gradiente como:

$$\frac{\partial \varepsilon(n)}{\partial w_{kj}(n)} = \frac{\partial \varepsilon(n)}{\partial e_k(n)} \frac{\partial e_k(n)}{\partial y_k(n)} \frac{\partial y_k(n)}{\partial v_k(n)} \frac{\partial v_k(n)}{\partial w_{kj}(n)} \quad (5)$$

Após manipulações algébricas obtém:

$$\frac{\partial \varepsilon(n)}{\partial w_{kj}(n)} = -e_k(n) \varphi_k'(v_k(n)) y_j(n) \quad (6)$$

A correção  $\Delta w_{kj}(n)$  aplicada a  $w_{kj}(n)$  é definida pela regra delta:

$$\Delta w_{kj}(n) = -\eta \frac{\partial \varepsilon(n)}{\partial w_{kj}(n)} \quad (7)$$

onde  $\eta$  é o parâmetro da taxa de aprendizagem do algoritmo de retropropagação. O uso do sinal negativo em (6) indica a descida do gradiente no espaço de pesos, ou seja, busca uma direção para a mudança de peso que reduz o valor de  $\varepsilon(n)$ .

Finalmente, temos a atualização dos pesos realizada por:

$$w_{kj}(t+1) = w_{kj}(t) - \eta \frac{\partial \varepsilon}{\partial w_{kj}}(t) \quad (8)$$

onde  $\omega_{kj}$  é o valor do peso da conexão  $k$  do neurônio  $j$  a ser atualizado.

### 2.3.2. Algoritmo Resilient Propagation (Rprop)

O algoritmo *Rprop*[10] apresenta um eficiente esquema de aprendizagem, tem uma adaptação direta do peso baseada na informação do gradiente local. Diferente do algoritmo do gradiente descendente com momento (*GDM*) a adaptação do *Rprop* não é interferida pelo comportamento da magnitude do gradiente da função erro.

Uma boa escolha dos parâmetros a fim de garantir a convergência deste método não é necessária sendo esta uma de suas principais vantagens.

No processo de otimização deste algoritmo cada peso será alterado individualmente através de um valor adaptativo  $\Delta_{ij}$ . Durante o processo de aprendizagem,  $\Delta_{ij}$  é baseado na função do erro de acordo com a seguinte regra de aprendizagem:

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^+ \Delta_{ij}^{(t-1)}, & \text{se } \frac{\partial E^{(t-1)}}{\partial \omega_{ij}} \frac{\partial E^{(t)}}{\partial \omega_{ij}} > 0 \\ \eta^- \Delta_{ij}^{(t-1)}, & \text{se } \frac{\partial E^{(t-1)}}{\partial \omega_{ij}} \frac{\partial E^{(t)}}{\partial \omega_{ij}} < 0 \\ \Delta_{ij}^{(t-1)}, & \text{caso contrário} \end{cases} \quad (9)$$

onde  $0 < \eta^- < 1 < \eta^+$

Podemos interpretar da seguinte forma: quando a derivada parcial da função erro em relação ao peso  $\omega_{ij}$  muda de sinal indica a última atualização foi muito grande, fazendo o erro da função ultrapassar o mínimo local, então o valor  $\Delta_{ij}$  é decrementado pelo fator  $\eta^- \Delta_{ij}$ . No caso da derivada manter seu sinal  $\Delta_{ij}$  permanecerá o mesmo com o objetivo de acelerar o processo de convergência.

Desta forma os incrementos (positivos ou negativos) dos pesos serão alterados conforme a equação (10).

$$\Delta \omega_{ij}^{(t)} = \begin{cases} -\Delta_{ij}, & \text{se } \frac{\partial E^{(t)}}{\partial \omega_{ij}} > 0 \\ +\Delta_{ij}, & \text{se } \frac{\partial E^{(t)}}{\partial \omega_{ij}} < 0 \\ 0, & \text{caso contrário} \end{cases} \quad (10)$$

A atualização dos pesos é realizada conforme a equação ( 11 )

$$\omega_{ij}^{(t+1)} = \omega_{ij}^{(t)} + \Delta \omega_{ij}^{(t)} \quad (11)$$

Podemos observar que quando a derivada é positiva o peso é diminuído do seu valor de atualização, caso contrário, o valor será adicionado.

Nos casos em que a derivada parcial muda de sinal, ou seja, o passo previsto foi muito grande passando do mínimo desejado, a atualização do peso deve ser realizada de acordo com a equação ( 12 ).

$$\omega_{ij}^{(t)} = -\Delta \omega_{ij}^{(t-1)}, \text{ se } \frac{\partial E^{(t-1)}}{\partial \omega_{ij}} \frac{\partial E^{(t)}}{\partial \omega_{ij}} < 0 \quad (12)$$

Espera-se que no próximo passo essa derivada mude novamente seu sinal. Para que não haja uma segunda “punição”, então se deve prever uma maneira de não fazer a adaptação do peso. Na prática, isto pode ser conseguido fazendo  $\frac{\partial E^{(t-1)}}{\partial \omega_{ij}} = 0$  na regra de adaptação de  $\Delta_{ij}$ .

### 2.3.3. Algoritmo Levenberg-Marquardt

Esse algoritmo é considerado o método mais rápido para treinamento de redes *feedforward backpropagation*, desde que a rede possua uma quantidade moderada de pesos sinápticos a fim de tornar o problema exequível. Para acelerar o treinamento da rede este algoritmo baseia-se na determinação das derivadas de segunda ordem do erro quadrático em relação aos pesos, diferindo do algoritmo backpropagation tradicional que considera as derivadas de primeira ordem.

O algoritmo de *Levenberg - Marquardt* ([11], [12]e[13]) se baseia no método de otimização de Newton, que faz uso da matriz Hessiana ( $H$ ). No método de *Levenberg -*

*Marquardt* se faz uma aproximação para essa matriz, mostrada na equação ( 13 ), determinada em função da matriz Jacobiana ( $J$ ), que contém as primeiras derivadas dos pesos em função dos pesos sinápticos, expressa na equação ( 14 ).

$$H = \frac{\partial^2 E(\omega)}{\partial \omega^2} \quad (13)$$

$$J = \frac{\partial E(\omega)}{\partial \omega} \quad (14)$$

onde  $E(\omega)$  é definido conforme a expressão ( 15 )

$$E(\omega) = \sum_{i=1}^n (y_i - y_{ei}) \quad (15)$$

A determinação da matriz Jacobiana é muito mais simples que a determinação da matriz Hessiana. Como, para uma rede neural, o desempenho de treinamento é expresso em função da soma dos erros quadráticos, a matriz Hessiana pode ser expressa pela expressão em ( 16 ).

$$H = J^T(\omega) \cdot J(\omega) \quad (16)$$

O método de Newton atualiza os pesos segundo a equação ( 17 ).

$$\omega_{k+1} = \omega_k - H^{-1} q_k \quad (17)$$

onde  $q_k$  pode ser escrito conforme ( 18 ).

$$q_k = 2 \cdot J^T(\omega) \cdot e(\omega) \quad (18)$$

O algoritmo de *Levenberg -Marquardt* procede a atualização dos pesos baseado na mesma expressão do método de Newton, realizando as modificações para a determinação da matriz Hessiana, mostrada em ( 19 ).

$$\omega_{k+1} = \omega_k - [J^T(\omega) \cdot J(\omega) + \mu_k I]^{-1} \cdot J^T(\omega) \cdot E(\omega) \quad (19)$$

onde,  $I$  é a matriz identidade e  $\mu_k$  é a constante do método de *Levenberg -Marquardt*.

O parâmetro  $\mu_k$  funciona como um fator de estabilização do treinamento, ajustando a aproximação de forma a utilizar a rápida convergência do método de Newton e evitando passos muito grandes que possam levar a um erro de convergência.

Esse método apresenta convergência em menos iterações, mas requer mais cálculos por iteração devido ao cálculo de matrizes inversas. Apesar do grande esforço computacional, ele segue sendo o algoritmo de treinamento mais rápido para redes neurais, quando se trabalha com um número moderado de parâmetros na rede. Se esse número é elevado, a utilização desse algoritmo é pouco prática.

### 3. Teoria dos Conjuntos Fuzzy

A teoria dos conjuntos fuzzy (nebulosos) é uma técnica inteligente que associa os conceitos básicos de funções de pertinência da teoria de conjuntos clássicos. Ela foi desenvolvida por Lofti A. Zadeh em 1965 com o objetivo de modelar o modo aproximado de raciocínio, imitando a habilidade humana de tomar decisões em um ambiente incerto e sem precisão. A sua utilização é apropriada em sistemas que se caracterizam pela generalidade, ambiguidade e imprecisão para a extração de informações vagas. A lógica fuzzy pode ser definida como a parte da lógica matemática dedicada aos princípios formais do raciocínio incerto ou aproximado, portanto mais próxima do pensamento humano e da linguagem natural. Devido a esta propriedade e à capacidade de realizar inferências, a lógica fuzzy tem encontrado grandes aplicações em várias áreas como modelagem de sistemas e reconhecimento de padrões[14].

Enquanto os conjuntos tratados pela lógica booleana permitem apenas o uso de funções de associação binária, que identificam os elementos pertinentes ou não pertinentes

a uma classe, a lógica fuzzy admite, também, a possibilidade de uma associação parcial, isto é, elementos que são parcialmente pertinentes a uma ou mais classes.

Em conjuntos clássicos, apenas dois valores para  $\mu_A(x)$  são permitidos: o elemento pertence (verdadeiro = 1) ou não pertence a um determinado conjunto (falso = 0). Na teoria dos conjuntos fuzzy, a transição entre pertencer e não pertencer é gradual, ou seja, é um número real variando de forma contínua no intervalo de zero a um [15].

Assim, dado um conjunto  $A$  em um universo  $X$ , os elementos deste universo simplesmente pertencem ou não pertencem àquele conjunto. Isto pode ser expresso pela função característica mostrada na equação ( 20 ).

$$f_A(x) = \begin{cases} 1, & \text{se e somente se } x \in A \\ 0, & \text{se e somente se } x \notin A \end{cases} \quad (20)$$

O conjunto fuzzy generaliza a função  $f_A$  de modo para que ela assuma um número infinito de valores no intervalo  $[0,1]$ . Assim, um conjunto  $A$  de valores fuzzy do universo  $X$  é definido pelos seguintes pares ordenados [16].

$$A = \{(x, \mu_A(x)) | x \in X\} \quad (21)$$

Onde  $\mu_A(x)$  é uma função denominada nível de pertinência do ponto  $x$  ao conjunto  $A$ , que determina com que grau um objeto  $x$  pertence a um conjunto  $A$ .

Os Conjuntos Fuzzy possuem algumas características especiais, tais quais: corte  $\alpha$ , conjunto de Níveis, Suporte, altura e normalização. O corte  $\alpha$  ( ${}^\alpha A$ ) de um conjunto fuzzy  $A$  corresponde ao conjunto tradicional que contém todos os elementos do conjunto universo  $X$  com grau de pertinência em  $A$  maior ou igual a  $\alpha$ , enquanto que o corte  $\alpha$  forte ( ${}^{\alpha+} A$ ) contém todos os elementos em um conjunto universo  $X$  com grau maior que  $\alpha$ , onde  $\alpha \in [0,1]$ . O corte  $\alpha$  e o corte  $\alpha$  forte são representados da seguinte forma:

$${}^\alpha A = \{x \in X | A(x) \geq \alpha\} \quad (22)$$

$${}^{\alpha+} A = \{x \in X | A(x) > \alpha\} \quad (23)$$

O Conjunto de Níveis ( $\Lambda$ ) de um Conjunto Fuzzy  $A$  corresponde a um conjunto que contém todos os valores  $\alpha \in [0,1]$  e que representam Cortes  $\alpha$  de  $A$  distintos. O Conjunto de Níveis do Conjunto Fuzzy  $A$  é representado formalmente na equação

$$\Lambda A = \{\alpha \mid A(x) = \alpha \text{ para algum } x \in X\} \quad (24)$$

O conjunto suporte de um conjunto fuzzy  $A$  é o conjunto de elementos no universo  $X$  para os quais  $\mu_A(x) > 0$ . Um conjunto fuzzy cujo suporte é um único ponto  $x'$  com  $\mu_A(x') = 1$  é chamado de conjunto unitário fuzzy ou singleton. Assim, um conjunto fuzzy também pode ser visto como o mapeamento do conjunto suporte no intervalo  $[0,1]$ , o que implica em expressar o conjunto fuzzy por sua função de pertinência [17][17]. O suporte de  $A$  é exatamente o mesmo que o corte  $\alpha$  forte de  $A$  para  $\alpha = 0$ . O suporte de um conjunto é representado como  $S(A)$  ou  $\text{supp}(A)$ .

A Altura ( $h$ ) de um conjunto fuzzy  $A$  corresponde ao seu maior grau de pertinência, entre todos os elementos do conjunto. Ela é representada da seguinte forma:

$$h(A) = \sup_{x \in X} A(x) \quad (25)$$

Um Conjunto Fuzzy  $A$  é chamado de Normal quando a sua Altura é igual a 1, ou seja, pelo menos um grau de pertinência, dos elementos do conjunto, possui valor máximo, enquanto que os conjuntos que não possuem Altura igual a um são chamados de subnormal. Portanto:

$$\begin{aligned} A \text{ é dito normal se } h(A) &= 1 \\ A \text{ é dito subnormal se } h(A) &< 1 \end{aligned} \quad (26)$$

Caso um Conjunto Fuzzy possua apenas um elemento com grau de pertinência igual a um, este elemento é denominado protótipo do conjunto. Um Conjunto Fuzzy não normalizado pode ser normalizado por meio da divisão dos graus de pertinência de cada elemento, pelo maior grau de pertinência encontrado no conjunto.

### 3.1. Propriedades e operações

As propriedades de união e intersecção fuzzy, bem como as demais, podem ser descritas através de operadores max e min. É fácil verificar que as seguintes propriedades algébricas de conjuntos ordinários também valem para conjuntos fuzzy:

Tabela 1 - Propriedades algébricas dos conjuntos

Involução	$(A')' = A$
Idempotência	$A \cap A = A$
	$A \cup A = A$
Comutatividade	$A \cap B = B \cap A$
	$A \cup B = B \cup A$
Associatividade	$(A \cap B) \cap C = A \cap (B \cap C)$
	$(A \cup B) \cup C = A \cup (B \cup C)$
Distributividade	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
Absorção	$A \cap (A \cup B) = A$
	$A \cup (A \cap B) = A$
Lei Transitiva	Se $A \subset B$ e $B \subset C \implies A \subset C$
Leis de Morgan	$(A \cap B)' = A' \cup B'$
	$(A \cup B)' = A' \cap B'$

Observando que as funções de pertinência dos conjuntos vazio ( $\emptyset$ ) e universo ( $X$ ) são definidas como sendo 0 e 1, respectivamente, verificam-se também as seguintes propriedades exibidas na equação (27).

$$\begin{aligned}
 (A \cap \emptyset) &= \emptyset & (A \cap X) &= A \\
 (A \cup \emptyset) &= A & (A \cup X) &= X
 \end{aligned}
 \tag{27}$$

Sabendo-se as propriedades do conjunto fuzzy, podem-se definir as principais operações e relações. Como citado anteriormente, a teoria dos conjuntos fuzzy é uma extensão da teoria dos conjuntos tradicionais, assim as operações e relações são definidas como extensão das operações e relações tradicionais, conforme Tabela 2, onde A e B são

conjuntos fuzzy sobre um conjunto base  $X$  e  $A(x)$  e  $B(x)$  representam os graus de pertinência de  $x$ .

**Tabela 2** – Operações e Relações com Conjuntos Fuzzy

Operação	Representação	Natureza
Complemento	$\neg A(x) = 1 - A(x)$	Operação
Diferença	$(A \neq B)$ se $A(x) \neq B(x)$ para pelo menos um elemento de $x \in X$	Relação
Igualdade	$(A = B)$ se $A(x) = B(x)$ para todo $x \in X$	Relação
Inclusão	$(A \subseteq B)$ se $A(x) \leq B(x)$ para todo $x \in X$	Relação
Intersecção	$A \cap B = A(x) \cap B(x) = \min [A(x), B(x)]$	Operação
União	$A \cup B = A(x) \cup B(x) = \max [A(x), B(x)]$	Operação

### 3.2. Variáveis linguísticas

Variável linguística é uma variável cujos valores são nomes de conjuntos fuzzy. Por exemplo, a estatura das pessoas pode ser uma variável linguística assumindo valores baixa, média, e alta. A estatura. Representa as funções de pertinências de acordo com os valores descritos do conjunto fuzzy[17].

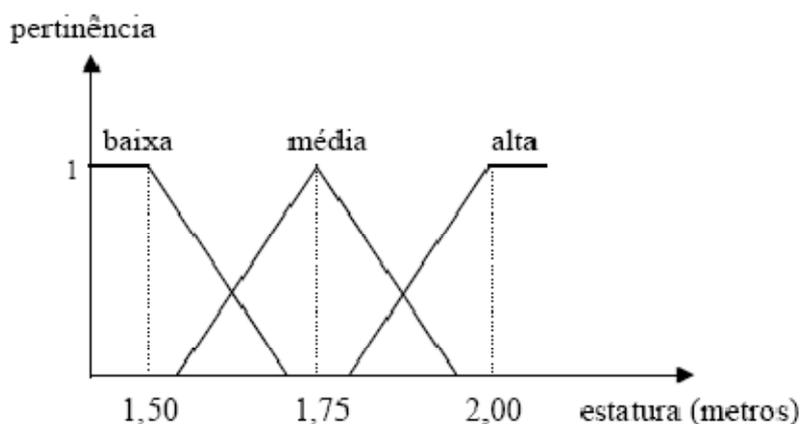


Figura 4- Funções de pertinência para a variável estatura.

Os valores de uma variável linguística podem ser formados a partir dos termos primários (alto, médio e baixo), de conectivos lógicos (e, ou e negação), de modificadores (muito, pouco, levemente e extremamente), e de delimitadores. Formalmente, uma variável linguística é caracterizada por uma quintupla  $(N, T(N), X, G, M)$ [17], onde:

N é o nome da variável;

T(N) é o conjunto de termos de N, ou seja, o conjunto de nomes dos valores linguísticos de N;

X é o universo de discurso;

G é a regra sintática para gerar os **valores** de N como uma composição de termos de T(N), conectivos lógicos, modificadores e delimitadores;

M é a regra semântica, para associar a cada valor gerado por G um conjunto fuzzy em X.

No exemplo dado acima, a variável linguística seria caracterizada da seguinte forma:

N: estatura;

T(N): {baixa, média, alta}

X: 0 a 2,00 m

G: estatura não baixa e não muito alta, por exemplo, M: associa o valor acima a um conjunto fuzzy cuja função de pertinência exprime o seu significado.

### 3.3. Funções de pertinência

As funções de pertinência podem ter diferentes formas, em geral, seu formato é restrito a certa classe de funções, representadas por alguns parâmetros específicos. Os formatos padrões mais comuns são de forma triangular, trapezoidal e Gaussiana.

a) Função triangular: parâmetros (a, m, b), com  $a \leq m \leq b$ .

$$\mu_A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x - a}{m - a} & \text{se } a < x \leq m \\ \frac{b - x}{b - m} & \text{se } m < x \leq b \\ 0 & \text{se } x > b \end{cases} \quad (28)$$

b) Função trapezoidal: parâmetros (a, m, n, b), com  $a \leq m$ ,  $n \leq b$  e  $m < n$ .

$$\mu_A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{m-a} & \text{se } a < x \leq m \\ 1 & \text{se } m < x \leq n \\ \frac{b-x}{b-n} & \text{se } n < x \leq b \\ 0 & \text{se } x > b \end{cases} \quad (29)$$

c) Função Gaussiana: parâmetros  $(m, \sigma_k)$ , com  $\sigma_k > 0$ .

$$\mu_A(x) = \exp[-\sigma_k(x-m)^2] \quad (30)$$

d) Singleton (Conjunto unitário): parâmetros  $(m, h)$ .

$$\mu_A(x) = \begin{cases} h & \text{se } x = m \\ 0 & \text{caso contrário.} \end{cases} \quad (31)$$

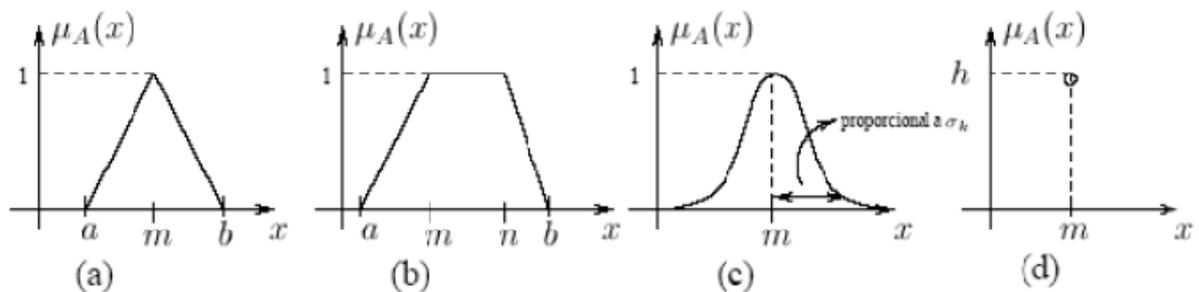


Figura 5 - Diferentes formatos das funções de pertinência - triangular (a), trapezoidal (b), Gaussiana (c), conjunto unitário (d)

As funções de pertinência contínuas podem ser definidas por intermédio de funções analíticas. Por exemplo, a seguinte função geral pode ser usada para definir as funções de pertinência associadas aos conjuntos fuzzy correspondentes aos termos pequeno, médio e grande [17]:

$$\mu_A(x) = (1 - (a(x-c))^b)^{-1} \quad (32)$$

Já as funções de pertinência descontínuas são compostas de segmentos contínuos lineares, resultando em formas triangulares ou trapezoidais. Funções de pertinência

discretizadas consistem de conjuntos de valores discretos correspondendo a elementos discretos do universo[17].

A escolha do formato mais adequado nem sempre é óbvia. A escolha de funções triangulares e trapezoidais é mais comum porque a ideia de se definirem regiões de pertinência total, média e nula é mais intuitiva do que a especificação do valor médio ( $m$ ) e dispersão ( $\sigma_k$ ), conceitos ligados às funções Gaussianas. Entretanto, as funções Gaussianas são bastante utilizadas devido às suas propriedades matemáticas de continuidade e diferenciação de classes[16].

Cada uma dessas funções de pertinências pode ser representada conforme a Figura 6, São elas[18]:

- a) pimf: é uma curva Pi-Shape;
- b) gaussmf: é uma curva simples de distribuição Gaussiana;
- c) gauss2mf: é composta pela diferença entre duas curvas de distribuição Gaussiana;
- d) gbellmf: é uma curva “generalizedbell”;
- e) dsigmf: é a diferença entre duas curvas sigmóides;
- f) psigmf: é o produto entre duas curvas sigmóides;
- g) trimf: é uma função triangular;
- h) trapmf: é uma função trapezoidal.

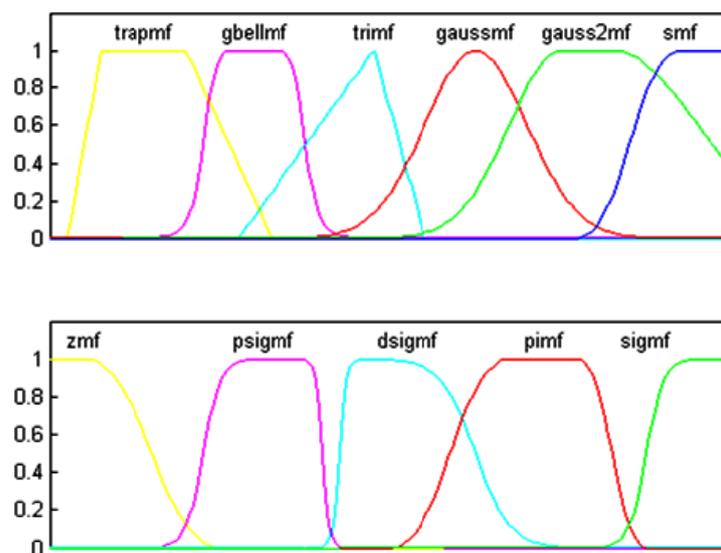


Figura 6–Funções de pertinência

### 3.4. Sistema de Inferência Mamdani

O modelo de inferência Mamdani foi um dos primeiros sistemas constituído utilizando a teoria de conjuntos fuzzy, sendo proposto em 1975 por Ebrahim Mamdani[20].

A regra de semântica tradicionalmente utilizada par ao processamento de inferências com o modelo de Mamdani é chamada de inferência Máx-Mín. Utilizando as operações de união e de interseção entre conjuntos da mesma forma de Zadeh[15].

As regras de produção em um modelo de Mamdani possuem relações fuzzy tanto em seus antecedentes como em seus consequentes.

Seja um sistema fuzzy composto de n regras, com uma das regras representada abaixo:

$$\text{Se } x_1 = A_1 \text{ e } x_2 = A_2 \text{ ... e } x_p = A_j \text{ então } y_1 = B_i,$$

onde,  $x_i$  são as entradas do sistema,  $A_1 \dots A_j$  são variáveis linguísticas definidas pelas funções de pertinência de entrada,  $y_1$  é a saída e  $B_i$  são as variáveis linguísticas definidas pelas funções de pertinência de saída.

O processo de inferência pode ser dividido em cinco etapas:

1ª etapa: Fuzzificação das Entradas

Nesta etapa toma-se o valor de cada variável de entrada e determina-se o seu grau de pertinência para cada uma das regras, ou seja:

$$\mu_{A_1}^k(x_1), \mu_{A_2}^k(x_2), \dots, \mu_{A_j}^k(x_j) \quad \text{onde } k = 1, \dots, n \quad (33)$$

2ª etapa: Aplicação do Operador fuzzy

Com as entradas fuzzificadas sabe-se qual o grau que cada parte do antecedente satisfaz cada regra. Precisa-se gerar o coeficiente de disparo de cada regra  $D^{(k)}$ , para isso aplica-se o operador fuzzy presentes no antecedente, no exemplo têm-se o operador “e”.

Para aplicar o operador o sistema de Mamdani utiliza a função mín, conforme equação abaixo:

$$D^{(k)} = T[\mu_{A_1}^k(x_1), \mu_{A_2}^k(x_2), \dots, \mu_{A_j}^k(x_j)] = \min[\mu_{A_1}^k(x_1), \mu_{A_2}^k(x_2), \dots, \mu_{A_j}^k(x_j)] \quad (34)$$

3ª etapa: Aplicação do método de implicação

O método de implicação é definido como a modelagem do conseqüente com base no coeficiente de disparo. A implicação se dá em todas as regras. O modelo de Mamdani utiliza a função *mim* que trunca a saída do conjunto fuzzy.

$$S^{(k)} = T[D^k, \mu_{B_i}(y)] = \min[D^k, \mu_{B_i}(y)] \quad (35)$$

4ª etapa: Agregação das Saídas

Ao realizar o processo de implicação cada regra gera-se uma função de pertinência truncada para a saída da regra. Como o sistema possui *n* regras devem-se ter *n* funções de pertinências truncadas que iram gerar a função de saída para isso agregam-se todos os gráficos de cada uma das funções, no modelo Mamdani utiliza-se a função *máx*.

$$\mu_{B'} = S[T[D^k, \mu_{B_i}(y)]] = \max[\min[D^k, \mu_{B_i}(y)]] \quad (36)$$

5ª etapa: Defuzzificação

Nesta etapa é realizada a conversão fuzzy-escalar, ou seja, transformam-se informações qualitativas em uma informação quantitativa.

Os métodos mais utilizados para realizar esta conversão são os métodos do centro de massa e o método da média dos máximos.

### 3.5. Sistema de Inferência Takagi-Sugeno

Na década de 80 um novo modelo de inferência baseado na teoria dos conjuntos fuzzy foi proposto. Denominado de modelo Sugeno, modelo de inferência fuzzy paramétrico ou simplesmente modelo TSK. As pesquisas mostraram que este modelo conseguia respostas satisfatórias para problemas que fossem representados satisfatoriamente apenas pelas suas relações entrada e saída.

Diferente do modelo de Mamdani o modelo de inferência do tipo TSK não utilizam funções de pertinência no conseqüente, as relações de saída são compostas de equações paramétricas que relacionam as entradas e saída do processo.

Seja um sistema fuzzy composto de *n* regras para o modelo TSK uma das regras seria representada conforme abaixo:

$$\text{Se } x_1 = A_1 \text{ e } x_2 = A_2 \text{ ... e } x_p = A_p \text{ então } y_1 = \phi(x_1, x_2, \dots, x_p)$$

O processo de inferência do modelo TSK é similar ao modelo Mamdani com algumas alterações. Na 2ª etapa do processo onde se calcula o valor do grau de disparo a função do operador normalmente não é a função mín. Além de não existir etapa de defuzzificação.

A saída do modelo TSK é calculada como a média ponderada das saídas de cada uma das regras que compõem o sistema, onde os pesos são os coeficientes de disparo.

Seja o coeficiente de disparo de cada regra dado por:

$$D^{(k)} = T[\mu_{A_1}^k(x_1), \mu_{A_2}^k(x_2), \dots, \mu_{A_j}^k(x_j)] \quad \text{onde} \quad k = 1, \dots, n \quad (37)$$

A saída do sistema TSK será dada por:

$$y = \frac{\sum_{k=1..n} D^{(k)} \cdot \phi(x_1, \dots, x_p)}{\sum_{k=1..n} D^{(k)}} \quad (38)$$

A eficiência do modelo de TSK está diretamente ligada a escolha dos parâmetros da função paramétrica de saída, normalmente os índices desta função são estimados seguindo algum índice de desempenho definido pelo usuário. A minimização do erro quadrático entre a saída do modelo de Sugeno e os dados de saída disponíveis é normalmente utilizada como medida de desempenho.

## 4. Óleo Isolante

O óleo isolante desempenha uma função importante no isolamento e no resfriamento das partes ativas de um transformador e, para o perfeito funcionamento deste as características físico-químicas do óleo devem ser preservadas.

As propriedades do óleo isolante podem sofrer alterações devido à oxidação, umidade, temperatura, instabilidade elétrica e também devido ao contato com materiais utilizados na construção de transformadores como papel, verniz, madeira, ar, silício, papelão e etc.

Em condições normais de uso existe uma degradação lenta das propriedades do óleo isolante, contudo, quando o transformador é submetido a distúrbios elétricos e térmicos, o óleo apresenta níveis de degradação superiores gerando como consequência gases a uma taxa maior que o normal.

Os estudos sobre gases em óleo devido a estresses térmicos e elétricos datam da década de 30, estes estudos mostraram que sob estresse térmico e elétrico, as moléculas de hidrocarbonetos presentes no óleo isolante podem se decompor e formar fragmentos ativos de hidrogênio e hidrocarbonetos. Estes fragmentos se combinam e formam gases como o hidrogênio ( $H_2$ ), metano ( $CH_4$ ), etano ( $C_2H_6$ ), etileno ( $C_2H_4$ ), acetileno ( $C_2H_2$ ) dentre outros.

Na década de 70 modelos termodinâmicos foram propostos para descrever a relação entre o estresse de temperatura e as características dos gases e de acordo com este modelo, a taxa de crescimento de cada gás produzido pode ser calculada para qualquer temperatura do óleo isolante. A figura abaixo ilustra a relação entre a temperatura do óleo e os gases gerados.

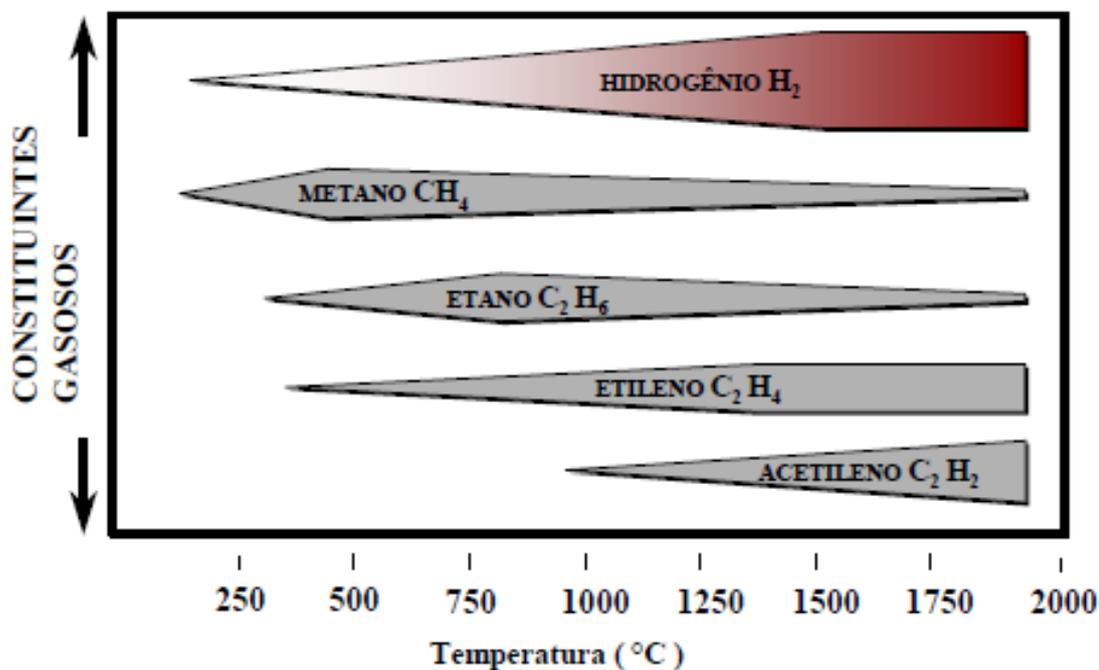


Figura 7 – Relação entre Temperatura e Geração de Gases no Óleo Isolante[19].

A figura acima possibilita revelar que os gases hidrogênio, metano, etano, etileno e acetileno são gerados seguindo uma ordem crescente de temperatura. O hidrogênio é gerado em baixas temperaturas (100 °C) enquanto o acetileno é gerado em altas temperaturas (1000 °C).

#### 4.1. Análise dos gases dissolvidos no óleo isolante

A partir da análise de geração dos gases em função da temperatura é possível analisar estado de funcionamento do transformador partindo da concentração de gases contidos na análise cromatográfica do óleo isolante.

Como técnica preditiva para avaliação e diagnóstico do estado de transformadores imersos em óleo mineral isolante, destaca-se o ensaio de análises de gases dissolvidos (DGA) ou análise cromatográfica do óleo isolante, que em função do tipo e da quantidade de gases combustíveis gerados pelo equipamento e absorvidos pelo óleo isolante, indica a presença de falhas incipientes associadas à parte ativa do transformador.

Dentro dos diversos métodos de diagnósticos para as análises cromatográficas do óleo isolante destacamos:

- Método pela IEC
- Método Rogers

Existem diversos outros métodos, muitas vezes de propriedade exclusiva de uma empresa, que não serão abordados nesse trabalho por não serem de uso tão difundido e generalizado quanto os citados acima, os quais foram utilizados na composição de normas internacionalmente reconhecidas.

Esses métodos são utilizados pela maioria das empresas do setor elétrico e fabricantes de equipamentos, sendo de conhecimento comum e de validade comprovada em todo o mundo. Caracterizados por experiências, muitas vezes empíricas, que associam as formações de alguns gases chaves, ou seja, o gás mais significativo gerado, além do tipo de falha incipiente ser associado a relações dos gases combustíveis presentes nas análises.

Os diagnósticos são obtidos a partir da distribuição das relações dos gases combustíveis em faixas, onde cada faixa é associada a um código numérico e cada código indica diferentes tipos de problemas internos que já são tabelados e mais prováveis de existirem e de serem responsáveis por esta geração de gases.

Atualmente a CELPE realiza esses ensaios de análises de gases dissolvidos (DGA). Dentro dessa análise, o critério vastamente utilizado pelas empresas contratadas pela CELPE para análise do óleo isolante é o critério de diagnóstico pelo método da IEC.

Conforme a literatura existente sobre o assunto, pode-se verificar que os diagnósticos realizados pelo método da IEC possuem um maior êxito, seguido do método de Rogers. O banco de dados da CELPE foi analisado e são descritas as metodologias de diagnóstico pelos métodos IEC e Rogers para famílias de transformadores.

A análise cromatográfica é feita a partir de uma amostra de óleo coletada de um transformador energizado. O objetivo deste tipo de análise é identificar e quantificar a presença dos principais gases dissolvidos no óleo, sendo os mais indicativos listados abaixo.

- Hidrogênio (H<sub>2</sub>)
- Nitrogênio (N<sub>2</sub>)
- Oxigênio (O<sub>2</sub>)
- Monóxido de carbono (CO)
- Metano (CH<sub>4</sub>)
- Dióxido de Carbono (CO<sub>2</sub>)
- Etileno (C<sub>2</sub>H<sub>4</sub>)
- Etano (C<sub>2</sub>H<sub>6</sub>)

- Acetileno ( $C_2H_2$ )

Dentre estes gases formados durante a operação do transformador os que são considerados não combustíveis são:

- Dióxido de Carbono ( $CO_2$ )
- Nitrogênio ( $N_2$ )
- Oxigênio ( $O_2$ ).

Os gases que se formam durante a operação normal do transformador são:

- Em maior concentração:
  - Monóxido de carbono ( $CO$ )
  - Dióxido de carbono ( $CO_2$ )
- Em pequenas concentrações:
  - Hidrogênio ( $H_2$ )
  - Metano ( $CH_4$ )
  - Etano ( $C_2H_6$ ).

Tradicionalmente esses ensaios são realizados a partir da retirada de uma amostra do óleo isolante do transformador e enviados ao laboratório, que irá medir a concentração dos gases encontrados no óleo e fornecer um diagnóstico baseado em normas e métodos conhecidos internacionalmente. A amostragem do óleo é feita normalmente com periodicidade anual, porém durante esse intervalo poderão surgir defeitos incipientes que não serão detectados por outras técnicas e levar o equipamento a operar em uma situação perigosa.

Quando a celulose é sobreaquecida ( $140^\circ C$ ) formam-se, além da água, os gases  $CO$ ,  $CO_2$ , que indicam a degradação celulósica da isolação. Já quando o óleo mineral isolante é sobreaquecido ( $500^\circ C$ ), originam-se os gases  $CH_4$ ,  $C_2H_6$ ,  $C_2H_4$ ,  $CO_2$  e água em presença de oxigênio. A pirólise do óleo, ou seja, a destruição do mesmo pelo calor conduz a formação de hidrogênio em maior quantidade, acetileno e em quantidade muito menor o metano e etileno.

Além do aquecimento, existem vários problemas que podem ocorrer no transformador, em maior ou menor intensidade, e que aceleram a formação de gases, acarretando problemas como as descargas parciais (efeito corona), o centelhamento

(descargas elétricas fracas com durações muito curtas), o arco elétrico, o sobreaquecimento localizado (ponto quente ou não), as correntes de circulação entre núcleo e tanque do equipamento e a hidrólise devido a presença de água.

Também pode ocorrer a geração de gases no óleo isolante do tanque principal o que, apesar de não constituir uma falha do transformador, indica uma situação irregular do mesmo. Esses gases gerados são provenientes de diferentes problemas assim como a contaminação do gás nitrogênio de transformadores selados, transformador não desgaseificado após ter sofrido reparos internos, projeto deficiente aonde exista contaminação dos gases gerados normalmente no comutador (CDC), pela chave comutadora, para o óleo do tanque principal do equipamento, equipamento transportado com CO<sub>2</sub>, aonde os enrolamentos absorvem parte desse gás, tratamento inadequado do óleo a ser utilizado no equipamento, por não ter desgaseificado corretamente este óleo e ainda a coleta inadequada.

A forma correta de realizar a coleta do óleo isolante não será tema deste trabalho, haja vista que já existem procedimentos normatizados que devem ser atendidos para a realização dessa tarefa.

A análise cromatográfica permite conhecer a porcentagem de cada gás existente no óleo isolante, e a partir destas concentrações, detectar um defeito no transformador. Os métodos da IEC e o de Rogers são os únicos que além de indicar os prováveis defeitos ocorrendo no equipamento, também indicam o seu funcionamento normal.

#### **4.1.1. Método IEC**

O método pela IEC (International Electrotechnical Commission) foi apresentado em 1978 em sua primeira edição da norma IEC, e atualizado em sua segunda edição em março de 1999 pela norma IEC inclusive com poucas alterações, gerando o método IEC revisada ou IEC2.

Este método é o mais utilizado, pois apresenta maior eficiência e melhor definição para um diagnóstico coerente a partir de análises cromatográficas do óleo isolante dos transformadores. Este método utiliza a concentração de cinco gases, metano (CH<sub>4</sub>), hidrogênio (H<sub>2</sub>), etano (C<sub>2</sub>H<sub>6</sub>), etileno (C<sub>2</sub>H<sub>4</sub>) e o acetileno (C<sub>2</sub>H<sub>2</sub>), utilizando as seguintes relações C<sub>2</sub>H<sub>2</sub>/C<sub>2</sub>H<sub>4</sub>, CH<sub>4</sub>/H<sub>2</sub>, e C<sub>2</sub>H<sub>4</sub>/C<sub>2</sub>H<sub>6</sub> para estabelecer os seus diagnósticos.

Os valores dessas relações são distribuídos por faixas e cada faixa corresponde a um código. Por sua vez, esses códigos geram algumas combinações que indicam os diagnósticos tabelados.

Os diagnósticos sugeridos pelo método IEC e que continuaram em sua revisão foram:

- Sem falhas ou envelhecimento normal.
- Descargas parciais de pequena densidade de energia
- Descargas parciais de alta densidade de energia
- Arco ou descargas de baixa energia
- Arco ou descargas de alta energia
- Sobreaquecimento ou defeito térmico de baixa temperatura (inferior a 150°C)
- Sobreaquecimento ou defeito térmico de baixa temperatura (150°C a 300°C)  
sobreaquecimento ou defeito térmico de média temperatura (300°C a 700°C)
- Sobreaquecimento ou defeito térmico de alta temperatura (superior a 700°C)

Cada diagnóstico tem a sua própria codificação quanto à presença de determinados gases no óleo isolante e em alguns casos não existe um diagnóstico para as concentrações dos gases encontradas no equipamento. No presente trabalho estes casos onde não são fornecidos diagnósticos pelas normas pesquisadas são tratados como não tabelados e os modelos desenvolvidos foram utilizados para fornecer uma aproximação de diagnósticos para estes casos.

A caracterização de um problema ou não no transformador é dada tanto pela taxa de evolução dos gases em mililitros por dia (ml/dia) como também por um valor limite de referência em ppm.

O grau de degradação da celulose é baseado na relação  $CO_2/CO$  que quando inferior a 3 indica uma situação considerada ruim. Entretanto para a confirmação deste laudo, devem-se realizar análises complementares através do grau de polimerização do papel e da medição do teor de furfuraldeído (2-Fal) no óleo isolante do transformador.

Uma contaminação do comutador CDC com o óleo isolante do tanque principal pode acontecer acarretando problemas na análise, pois mascara as reais condições de operação do transformador. Para que isso não ocorra, toda vez que à relação

$C_2H_2/H_2$  estiver maior que 2 ou 3 o que indica uma provável contaminação oriunda do comutador CDC, deve-se também realizar uma inspeção no comutador CDC/transformador para confirmar este diagnóstico.

A codificação pela da norma IEC 1 referente às relações dos gases é dada na Tabela 3.

**Tabela 3** – Codificação das relações dos gases pelo método da IEC

	Valor	Código
$R_1 = \frac{C_2H_2}{C_2H_4}$	$R_1 < 0.1$	0
	$0.1 \leq R_1 \leq 3$	1
	$R_1 > 3$	2
$R_2 = \frac{CH_4}{H_2}$	$R_2 < 0.1$	1
	$0.1 \leq R_2 \leq 1$	0
	$1 < R_2$	2
$R_3 = \frac{C_2H_4}{C_2H_6}$	$R_3 < 1$	0
	$1 \leq R_3 \leq 3$	1
	$R_3 > 3$	2

A Tabela 3 mostra os códigos gerados partindo das faixas de valores das relações das concentrações dos gases para o método IEC, com esses códigos são construídos os diagnósticos com as faltas incipientes que podem está ocorrendo no transformador.

Os diagnósticos tabelados que são sugeridos na norma IEC podem ser vistos na Tabela 4.

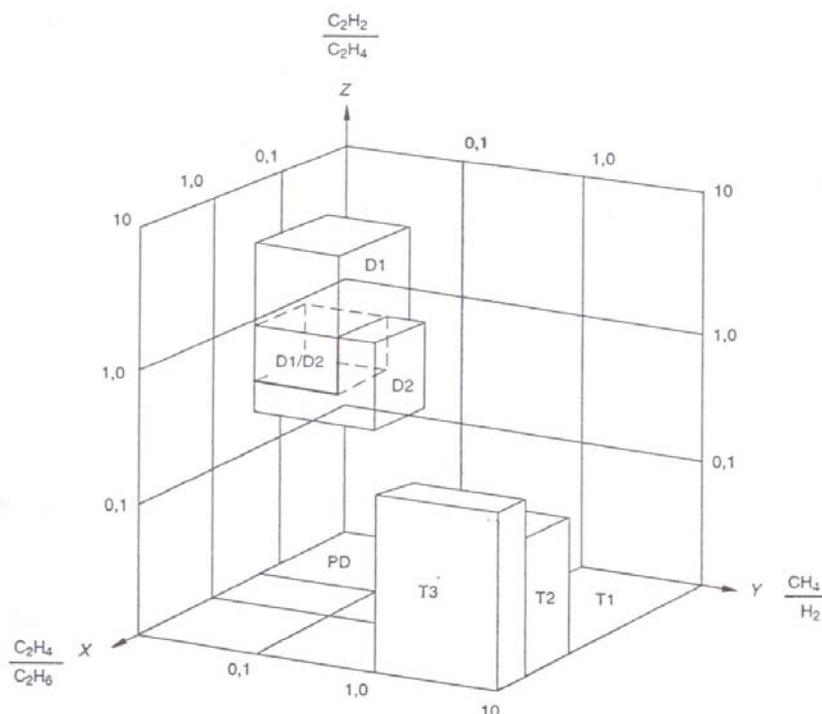
**Tabela 4** – Diagnósticos sugeridos pelo método da IEC

Código gerado	Diagnóstico sugerido
000	Sem Falha
001	Sobreaquecimento ou Defeito térmico de baixa temperatura (<150°C)
010	Descargas parciais de pequena densidade de energia
020	Sobreaquecimento ou Defeito térmico de baixa temperatura (150°C a 300°C)
021	Sobreaquecimento ou Defeito térmico de média temperatura (300°C a 700°C)
022	Sobreaquecimento ou defeito térmico de alta temperatura (superior a 700°C)
101	Arco ou Descargas de baixa energia
102	Arco ou Descargas de alta energia

110	Descargas parciais de alta densidade de energia
202	Arco ou Descargas de baixa energia

Essa codificação dos diagnósticos foi utilizada como dado de entrada dos protótipos criados durante a realização desse trabalho, como será mostrado em capítulo posterior.

A revisão da norma que traz o método IEC sugere uma nova maneira de visualizar a proximidade dos diagnósticos não tabelados com os tabelados, como pode ser observado no gráfico da Figura 8. Assim, pode-se aproximar um problema não tabelado por um defeito conhecido de acordo com a relação dos gases. Além disso, esta representação gráfica é conveniente, sobretudo, para o acompanhamento da evolução das falhas incipientes visualmente.



**Figura 8**– Representação gráfica das razões dos gases [3]

Onde : PD – Descargas parciais ou corona;

D1 – Descargas de baixa densidade de energia;

D2 – Descargas de alta densidade de energia;

T1 – Falta térmica,  $t < 300$  °C;

T2 – Falta térmica,  $300 < t < 700$  °C;

T3 – Falta térmica,  $t > 700\text{ }^{\circ}\text{C}$

#### 4.1.2. Método de ROGERS

Depois do método da IEC, o método de Rogers é o mais utilizado pelas empresas contratadas pela CELPE para realização das análises cromatográficas, além disso, em casos onde o método IEC não fornece um diagnóstico (casos não tabelados), o método ROGERS pode ser utilizado como indicativo de uma falha ocorrendo no transformador.

Este método utiliza quatro relações e cinco gases para estabelecer as condições do equipamento, sendo os gases o metano ( $\text{CH}_4$ ), hidrogênio ( $\text{H}_2$ ), etano ( $\text{C}_2\text{H}_6$ ), etileno ( $\text{C}_2\text{H}_4$ ) e o acetileno ( $\text{C}_2\text{H}_2$ ) e as relações utilizadas  $\text{CH}_4/\text{H}_2$ ,  $\text{C}_2\text{H}_6/\text{CH}_4$ ,  $\text{C}_2\text{H}_4/\text{C}_2\text{H}_6$ , e  $\text{C}_2\text{H}_2/\text{C}_2\text{H}_4$ , além dessas relações são também utilizadas as relações de  $\text{CO}_2/\text{CO}$  e  $\text{CO}/\text{TCG}$ , onde TCG significa total concentração dos gases combustíveis e é obtido a partir da soma das concentrações de todos os gases combustíveis, para avaliar se existe indicativo de sobreaquecimento da celulose.

A relação  $\text{CO}_2/\text{CO}$  é utilizada para avaliar a degradação da celulose e é sempre avaliada em conjunto com a relação do  $\text{CO}/\text{TCG}$ , a Tabela 5 mostra quais os possíveis diagnósticos obtidos a partir dessas relações.

**Tabela 5** – Avaliações da relação  $\text{CO}_2/\text{CO}$

<b><math>\text{CO}_2/\text{CO}</math></b>	<b><math>\text{CO}/\text{TCG}</math></b>	<b>Resultado</b>
<7 ou >11	>80%	Sobreaquecimento generalizado na celulose. Temperaturas acima do normal nas proximidades do isolamento celulósico.
<7 ou >11	< 80%	

Os defeitos de alta densidade de energia são percebidos através dos altos percentuais de hidrogênio e acetileno, com baixa quantidade de metano e etileno. Se a celulose também está envolvida a concentração de dióxido e monóxido de carbono é relevante. O total de gases gerados é muito alto, e deve-se levar em consideração que o acetileno por ser muito solúvel e reativo desaparece rapidamente, devido a sua combinação com outros gases.

Nas descargas com baixa densidade de energia, os primeiros gases a serem produzidos são o hidrogênio e, em menor quantidade, o metano. A produção desses gases se realiza em pequenas quantidades, entretanto podem se estender por um longo período antes do transformador falhar.

O sobreaquecimento local é identificado devido às significativas quantidades de hidrogênio, metano e etileno e em alguns casos pequenas quantidades de acetileno. Se grandes quantidades de dióxido de carbono e baixas de monóxido de carbono são encontradas, então a celulose também está envolvida.

Descargas de corona são caracterizadas pela formação predominantemente de hidrogênio e significativa quantidade de metano, com traços de etano e etileno.

A decomposição térmica da celulose é o resultado do sobreaquecimento da isolação do transformador devido à sobrecarga, ou onde a normal circulação do óleo é deficiente. Os gases produzidos são primariamente dióxido de carbono e monóxido de carbono com algum hidrogênio, ocasionalmente. Em alguns casos, o baixo peso molecular dos hidrocarbonetos pode resultar em colapso térmico do óleo associado.

Os diagnósticos fornecidos pelo método são obtidos a partir da construção de duas tabelas em que, na primeira (Tabela 6) os valores da divisão das concentrações dos gases, são distribuídos em faixas e cada relação recebe um valor numérico dependendo da faixa em que está incluso.

**Tabela 6** – Construção método ROGERS

Relação dos gases	Faixa de valores	Código numérico
$R_1 = \frac{CH_4}{H_2}$	$\leq 0,1$	5
	$]0,1 - 1[$	0
	$]1 - 3[$	1
	$\geq 3$	2
$R_2 = \frac{C_2H_6}{CH_4}$	$< 1$	0
	$\geq 1$	1
$R_3 = \frac{C_2H_4}{C_2H_6}$	$< 1$	0
	$]1 - 3[$	1
	$\geq 3$	2
$R_4 = \frac{C_2H_2}{C_2H_4}$	$< 0,5$	0
	$]0,5 - 3[$	1
	$\geq 3$	2

Na segunda tabela os códigos numéricos atribuídos às relações são agrupados, na ordem correta das relações e cada novo código formado pela junção dos demais exibe um diagnóstico que pode ser tabelado pelo método ou não. Na Tabela 7 são exibidos os códigos que são tabelados pelo método.

**Tabela 7** – Diagnósticos sugeridos

$R_1$	$R_2$	$R_3$	$R_4$	<b>Diagnóstico proposto</b>
0	0	0	0	Deterioração Normal
0	0	0	1	Descarga contínua
0	0	1	0	Sobreaquecimento de condutores
0	0	1	1	Arco com alta energia
0	0	1	2	Arco com alta energia
0	0	2	1	Arco com alta energia
0	0	2	2	Descarga continua de baixa potência
0	1	0	0	Defeito térmico de média temperatura 200-300°C
1	0	0	0	Defeito térmico de baixa temperatura < 150°C
1	0	1	0	Corrente de Circulação No Enrolamento
1	0	2	0	Corrente de Circulação No Núcleo do Tanque, Sobreaquecimento Em Conexões.
1	1	0	0	Defeito Térmico de Media Temperatura 150-200°C
2	0	0	0	Defeito Térmico de Baixa Temperatura < 150°C
2	1	0	0	Defeito Térmico de Media Temperatura 150-200°C
5	0	0	0	Descargas parciais
5	0	0	1	Descarga parcial envolvendo o papel
5	0	0	2	Descarga parcial envolvendo o papel

O diagnóstico de um transformador é obtido com base na avaliação dos resultados apresentados por cada análise cromatográfica do equipamento e também na evolução da taxa de gases das últimas amostras. Um especialista analisando essas duas conformidades informa que o equipamento está ou não satisfatório. Desta forma o equipamento pode permanecer em operação ou ser reavaliado para confirmar o diagnóstico. Se após a nova análise cromatográfica algum defeito no transformador seja confirmado, opta-se pela retirada do mesmo de operação, encaminhando-o para uma inspeção interna e posterior reparo. Caso contrário, o transformador é mantido em operação.

## **5. Implementação Redes Neurais**

### **5.1. Conjunto de dados**

Para realização do trabalho de manutenção preditiva de transformadores de potência foram utilizadas algumas etapas do processo de descoberta do conhecimento para produzir de forma automática ou semiautomática um conjunto de padrões que poderão ser utilizados posteriormente por diferentes técnicas de Redes Neurais. Este processo envolveu várias etapas, entre elas:

- Digitalização das informações;
- Seleção de dados;
- Pré-Processamento dos dados;

Os dados utilizados para a concepção da pesquisa foram cedidos pela Companhia Elétrica de Pernambuco – CELPE e são obtidos a partir da realização de análises por empresas contratadas. Estes dados contidos nos laudos muitas vezes variavam de empresa para empresa e por isso foi necessário definir quais deles seriam relevantes para o estudo da manutenção preditiva de transformadores de potência.

Os dados fornecidos pela CELPE encontravam-se distribuídos em laudos impressos armazenados em arquivos. Assim, a primeira dificuldade encontrada foi à necessidade de digitalizar esta vasta gama de dados para uma planilha no Excel.

No total foram 4905 laudos digitalizados, cada um deles com seus respectivos campos relevantes para o arquivamento de informações e um melhor acompanhamento histórico das análises cromatográficas de gases dissolvidos em óleos isolantes dos transformadores da CELPE. Fica claro a grande dificuldade e o dispendioso tempo utilizado para realização desta tarefa no início do desenvolvimento do trabalho.

#### **5.1.1. Pré-processamento dos dados**

Muitas vezes os dados originais estão incompletos ou inconsistentes. Técnicas de pré-processamento dos dados podem aumentar a qualidade dos dados, por meio do aumento da exatidão. O pré-processamento dos dados é um passo importante no processo de descoberta do conhecimento, desde que decisões de qualidade se baseiam na qualidade

dos dados utilizados. Detectar anomalia nos dados corrigi-las o mais cedo possível e reduzir os dados a serem analisados, possibilitam uma tomada de decisão de qualidade.

### 5.1.2. Seleção de Atributos

Nesta etapa, os dados são consolidados na forma apropriada para a utilização dos mesmos pelas técnicas de Redes Neurais. A forma de representação dos dados de entrada e de saída para um problema de aprendizado de redes neurais é uma das decisões chave que influenciam na qualidade da solução que pode ser obtida[21].

. Dependendo do tipo do problema, existem várias maneiras diferentes de representar os atributos. Para todos os tipos de atributos, existem muitos métodos de representação. Os métodos que serão utilizados para representar os atributos são os métodos recomendados pelo relatório do *Proben* [21]e serão descritos a seguir.

- **Atributos reais**

Atributos com valores reais devem ser reescalados. O ideal é que os valores reais estejam compreendidos entre intervalos bem definidos. Valores em escalas diferentes podem confundir o aprendizado da rede neural, fazendo com que esta tenha mais consideração por valores com grandes magnitudes. A normalização se faz necessária para que os valores esparsos não interfiram no processo de aprendizado e convergência do algoritmo selecionado para treinamento das redes. Para redes neurais, a normalização também auxilia na efetividade e rapidez do treinamento.

Vários tipos de pré-processamento podem ser utilizados com o objetivo de normalizar os sinais em uma faixa de valores, dentre o quais poderíamos destacar a equação ( 39 )que realiza a normalização para que os valores fiquem no intervalo [0 1] e a equação ( 40 )que realiza a normalização para que os valores fiquem no intervalo[-1 1].

$$x_{normalizado} = \frac{x - x_{mínimo}}{x_{máximo} - x_{mínimo}} \quad ( 39 )$$

$$x_{normalizado} = \frac{(2x - x_{máximo} - x_{mínimo})}{x_{máximo} - x_{mínimo}} \quad (40)$$

- **Atributos inteiros**

Atributos com valores inteiros são frequentemente manipulados como se fossem valores reais. Se o número de diferentes valores é pequeno, uma das representações utilizadas para valores nominais também pode ser apropriada.

- **Atributos nominais**

Atributos com  $m$  diferentes valores são representados pela codificação *1-of-m* ou código binário. Para a codificação *1-of-m*, são criados  $m$  novos atributos de forma que cada vetor representa uma única classe, onde neste vetor de tamanho  $m$  apenas uma das posições apresenta valor 1 e as demais zero, como mostra o exemplo da Tabela 8.

**Tabela 8** - Codificação dos dias da semana

<b>Dia da semana</b>	<b>Código correspondente</b>
Domingo	1000000
Segunda-feira	0100000
Terça-feira	0010000
Quarta-feira	0001000
Quinta-feira	0000100
Sexta-feira	0000010
Sábado	0000001

Essa forma de codificação foi utilizada para representar as saídas da Rede Neural, pois os atributos de saída são os diagnósticos nominais fornecidos pelos métodos.

### **5.1.3. Elaboração do conjunto de dados para utilização na Rede Neural**

A seleção das características de entrada é um primeiro passo essencial à configuração de uma RNA. Os dados de entrada devem ser escolhidos de forma cuidadosa, pois estes refletirão as características do problema.

Para atender a estes requisitos todos os dados apresentados à rede neural foram pré-processados. O pré-processamento dos dados de entrada proporciona a RNA um aprendizado mais rápido e eficiente. Assim as entradas:  $X_1=C_2H_2/C_2H_4$ ;  $X_2=CH_4/H_2$ ; e  $X_3=C_2H_4/C_2H_6$  foram normalizadas conforme Equação ( 41 ).

$$X_i = \frac{|X_i - X_{\min}|}{|X_{\max} - X_{\min}|} \quad (41)$$

Onde:  $X_i = X_1, X_2$  ou  $X_3$  e  $X_{\min}$  é o menor valor encontrado no conjunto de  $X_i$ , dentro das 4905 amostras (tabeladas e não tabeladas). No banco de dados da CELPE, estes valores foram iguais a zero para  $X_{\min 1}$ ,  $X_{\min 2}$  e  $X_{\min 3}$ .  $X_{\max}$  é o valor máximo encontrado no conjunto de  $X_i$ . No banco de dados, estes valores foram:  $X_{\max 1} = 81,6$ ;  $X_{\max 2} = 261,25$  e  $X_{\max 3} = 330,0$ ,

Desses valores faziam parte alguns encontrados em amostras não tabeladas. Isso se fez necessário, pois as amostras não tabeladas foram simuladas na rede para encontrar uma aproximação de diagnóstico e para servirem de dado de entrada da rede também necessitam ser normalizadas.

Porém todas as amostras não tabeladas, aproximadamente 47% do total, foram extraídas dos conjuntos de treinamento, validação e teste das redes, para evitar uma distorção no aprendizado da rede já que essas amostras não apresentam resultados pelos métodos adotados. Ao longo do desenvolvimento do projeto, no entanto, foi possível simular algumas destas amostras e comparar o resultado fornecido pela rede com o resultado presente no laudo emitido pela empresa contratada para a análise do óleo isolante.

Dos 4905 laudos fornecidos pela empresa e digitalizados foram encontrados 2591 diagnósticos tabelados pelo método proposto na IEC e 2057 diagnósticos tabelados pelo método ROGERS, porém alguns dos diagnósticos possuíam uma quantidade muito pequena de amostras, como pode ser visto na Tabela 9 para o método IEC e na Tabela 10 para o método de ROGERS.

**Tabela 9** – Quantidade de diagnóstico por classe para o método IEC.

<b>Código IEC</b>	<b>Quantidade de amostras</b>	<b>Percentual dos tabelados</b>	<b>Diagnóstico</b>
000	571	22,0%	Sem falha

001	436	16,8%	Defeito térmico de baixa temperatura < 150 °C
010	237	9,1%	Descargas parciais de pequena densidade de energia
020	615	23,7%	Defeito térmico de baixa temperatura 150-300 °C
021	208	8,0%	Defeito térmico de media temperatura 300-700 °C
022	232	9,0%	Defeito térmico de alta temperatura > 700 °C
101	41	1,6%	Descargas de energia reduzida
102	143	5,5%	Descargas de alta energia
110	9	0,3%	Descargas parciais de alta densidade de energia
202	99	3,8%	Descargas de energia reduzida

A Tabela 9 exibe a quantidade de dados encontrados por diagnóstico sugerido pelo método IEC. É possível observar que existem diagnósticos com uma quantidade significativa de ocorrência, como o defeito térmico de baixa temperatura (código IEC 020) com 615 amostras que apontam para essa falha incipiente ocorrendo no transformador.

Contudo existem diagnósticos que são muito difíceis de ocorrer e mesmo com o longo período de amostragem que foi utilizado para a formulação do banco de dados existem apenas nove amostras da falha incipiente descargas parciais de alta densidade de energia (código IEC 110), correspondendo a menos de 0,5% do total de amostras tabeladas. O que tornou bastante difícil a utilização de ferramentas como Redes Neurais ou Anfis, que são ferramentas que necessitam de bastante dados para o seu treinamento e para tornar seu aprendizado generalizador.

**Tabela 10** - Quantidade de diagnóstico por classe para o método ROGERS.

<b>Código IEC</b>	<b>Quantidade de amostras</b>	<b>Percentual dos tabelados</b>	<b>Diagnóstico</b>
0000	200	9,7%	Deterioração normal
0001	13	0,6%	Descarga contínua
0010	276	13,4%	Sobreaquecimento de condutores
0011	16	0,8%	Arco com alta energia
0012	1	0,0%	Arco com alta energia
0021	72	3,5%	Arco com alta energia
0022	89	4,3%	Descarga contínua de baixa potencia
0100	371	18,0%	defeito térmico de media temperatura 200-300°C
1000	152	7,4%	Defeito térmico de baixa temperatura < 150°C
1010	131	6,4%	Corrente de circulação no enrolamento

1020	138	6,7%	Corrente de circulação no núcleo do tanque, sobreaquecimento em conexões
1100	167	8,1%	Defeito térmico de media temperatura 150-200°C
2000	133	6,5%	Defeito térmico de baixa temperatura < 150°C
2100	185	9,0%	Defeito térmico de media temperatura 150-200°C
5000	98	4,8%	Descargas parciais
5001	1	0,0%	Descarga parcial envolvendo o papel
5002	14	0,7%	Descarga parcial envolvendo o papel

Na Tabela 10 é indicada a quantidade de amostra por diagnóstico proposto pelo método de ROGERS, indicando também a quantidade percentual de diagnóstico pela quantidade total de diagnósticos tabelados.

Nota-se a mesma dificuldade encontrada quando trabalhando com o método IEC, já que algumas amostras possuem poucos diagnósticos para o treinamento das Redes Neurais, como a ocorrência de descarga parcial envolvendo o papel, que em todo o tempo de coleta de amostra do óleo só foi diagnosticada uma vez.

Para montagem dos conjuntos de dados para treinamento, validação e teste das RNAs foram utilizadas 100 amostras para cada classe de diagnóstico, mesmo nos casos onde havia dois códigos representando o mesmo defeito, como é o caso das Descargas de Energia Reduzida que tanto pode ser diagnosticado código 101 como pelo código 202, foram tratados como casos independentes para apresentação à rede e utilizou-se 100 amostras para cada código.

Nos casos onde os diagnósticos não possuíam uma quantidade de amostras suficientes para montar os conjuntos de treinamento, validação e teste, foram extraídos dados de outro banco que possuíam transformadores com as mesmas características, ou ainda, os dados tiveram de ser replicados para atingir a quantidade desejada.

Este procedimento faz-se necessário para evitar a tendência da rede em diagnosticar as classes com maior quantidade de padrões, evitando que a rede apenas “decore” as respostas e não generalize o problema. Sendo assim, o conjunto de dados de entrada da rede foi composto por 1000 amostras para o método IEC, sendo 100 de cada diagnóstico proposto, onde 50% dos dados foram utilizados para treinamento, 25% para validação e os outros 25% utilizados para o teste da rede.

Este conjunto de 1000 amostras foi selecionado de forma aleatória para compor os três conjuntos de treinamento, validação e teste, visando prevenir a tendência da rede de ajustar os pesos de acordo com a ordem de apresentação dos dados.

O conjunto de treinamento é utilizado para ensinar a rede por meio das alterações dos pesos das conexões das unidades de processamento; o conjunto de validação para verificar a eficiência da rede quanto a sua capacidade de generalização durante o treinamento, sendo utilizado geralmente como um dos critérios de parada do treinamento; e o de teste para verificar seu desempenho frente a dados reais nunca vistos anteriormente.

Vale salientar que o conjunto de teste, apesar de não ser composto das mesmas amostras presentes no seu treinamento, ele deve possuir as mesmas características intrínsecas das amostras utilizadas no treinamento da rede, no tópico seguinte será mostrada a arquitetura da rede escolhida.

## **5.2. Arquitetura da rede neural**

As RNAs foram designadas para diagnosticarem falhas em transformadores por causa das suas excelentes habilidades de aprender classificações complexas ([9] e [29]).

### **5.2.1. Rede Multilayer Perceptron – MLP**

As redes MLP são arranjos multicamadas de elementos de processamento não lineares que são constituídas por uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída. A rede *MLP* é um dos paradigmas mais famosos utilizados para implementar sistemas baseados em redes neurais. É capaz de solucionar problemas não lineares, coisa que as redes de uma camada não são capazes de resolver facilmente.

O processo de aprendizagem da rede *MLP* é dado da seguinte forma, um conjunto de padrões de entrada é apresentado à rede, que propaga estes sinais a cada camada, no sentido para frente (*forward*). A utilização do algoritmo de aprendizagem *backpropagation*, que consiste em comparar a saída desejada com a obtida pela rede e calcular o erro associado com cada padrão de entrada. Este erro é propagado, no sentido para trás (*backward*), da camada de saída para a camada de entrada, onde a rede será



de entrada correspondendo às razões dos gases  $C_2H_2/C_2H_4$ ,  $CH_4/H_2$  e  $C_2H_4/C_2H_6$ ; na camada de saída, possui 10 nodos, onde cada nodo de saída corresponde a um diagnóstico proposto pelo método IEC.

Desta forma, o vetor de entrada da RNA é dado por  $[C_2H_2/C_2H_4, CH_4/H_2, C_2H_4/C_2H_6]^T$ , e o de saída é um vetor com 10 dígitos onde apenas um elemento possui valor 1 (um), e os demais, valor 0 (zero).

Como em cada nodo de saída da rede são encontrados números reais e uma saída binária seria simulada, foi adotada a estratégia de “o vencedor leva tudo”, onde aquele nodo que possui o maior valor dentre todos os nodos recebe o binário “1” e os demais recebem o número binário “0”, assim são formados os códigos de saída referentes a cada diagnóstico, que podem ser vistos na Tabela 11.

**Tabela 11** – Código do vetor de saída da rede neural

<b>Código IEC</b>	<b>Código saída RNA</b>
000	1000000000
001	0100000000
010	0010000000
020	0001000000
021	0000100000
022	0000010000
101	0000001000
102	0000000100
110	0000000010
202	0000000001

### 5.2.2. Determinação do número de nodos da camada intermediária

Para a determinação do número de nodos da camada intermediária, alguns autores propõem determinadas regras heurísticas para se definir um modelo inicial para a arquitetura da rede a ser adotada. Uma dessas regras estipula que a camada intermediária deve ter no mínimo uma quantidade de nodos igual à média dos nodos utilizados nas camadas de entrada e saída da rede.

As redes foram treinadas com o algoritmo de treinamento Levenberg-Marquardt-LM[23]. Este algoritmo é mais rápido que outros algoritmos de treinamento de um fator que varia de 10 a 100 vezes. A desvantagem do algoritmo LM, é que ele requer muita memória no seu processamento. Porém, na versão 3 do MATLAB® [26], já houve uma

redução de memória no algoritmo, permitindo que este seja usado em problemas de maior dimensão.

Foram analisadas arquiteturas com 8, 9 e 10 nodos na camada intermediária, com o intuito de escolher a rede que apresentasse melhores resultados. A rede que foi desenvolvida com nove nodos na camada intermediária atingiu um desempenho satisfatório.

O critério adotado para a escolha da melhor arquitetura de rede foi a média do erro médio quadrado – MSE – no conjunto de validação da rede. Para minimizar o efeito das inicializações dos pesos, para cada quantidade de nodos foram criadas 10 redes com pesos iniciais aleatórios.

**Tabela 12** – Erro médio quadrado das redes.

	<b>8 Nodos</b>	<b>9 Nodos</b>	<b>10 Nodos</b>
<b>Rede 1</b>	2,01%	0,11%	2,10%
<b>Rede 2</b>	0,05%	1,18%	2,13%
<b>Rede 3</b>	0,39%	1,15%	2,10%
<b>Rede 4</b>	2,21%	1,07%	1,15%
<b>Rede 5</b>	0,06%	0,30%	0,84%
<b>Rede 6</b>	0,17%	2,71%	0,44%
<b>Rede 7</b>	1,46%	1,12%	3,03%
<b>Rede 8</b>	0,13%	1,16%	2,36%
<b>Rede 9</b>	3,14%	1,18%	1,19%
<b>Rede 10</b>	1,06%	0,15%	1,22%
<b>Média</b>	1,07%	1,01%	1,66%

A Tabela 12 exibe o percentual do MSE do conjunto de validação para as dez redes criadas mostrando que a rede com nove nodos na camada intermediária foi a que apresentou uma menor média dos erros médio quadrado e foi por isso a topologia escolhida.

Além do MSE do conjunto de validação as redes criadas foram ainda simuladas com o conjunto de teste, que possui dados ainda não apresentados à rede durante o treinamento, para confirmar o desempenho da rede apontada como melhor pelo critério do MSE do conjunto de validação.

A Tabela 13exibe o MSE percentual de classificação das redes no conjunto de teste em função do número de nodos na camada intermediária. Foram fornecidos 250 padrões (25 padrões por diagnóstico) para teste das redes.

O padrão escolhido com nove nodos além de apresentar a menor média dos MSE do conjunto de validação, também apresentou a menor média dos MSE do conjunto de teste durante a simulação e ainda a rede com o menor MSE no conjunto de teste, tendo a Rede 1 com nove nodos MSE=2%, como pode ser visto na Tabela 13.

**Tabela 13** – Erro percentual de classificação no conjunto de teste

	<b>8 Nodos</b>	<b>9 Nodos</b>	<b>10 Nodos</b>
<b>Rede 1</b>	22,4	2,0	24,0
<b>Rede 2</b>	4,4	15,6	25,6
<b>Rede 3</b>	5,6	12,8	22,0
<b>Rede 4</b>	24,0	14,0	13,2
<b>Rede 5</b>	4,4	4,8	6,4
<b>Rede 6</b>	5,2	25,2	6,0
<b>Rede 7</b>	18,4	13,2	33,2
<b>Rede 8</b>	4,4	13,6	25,2
<b>Rede 9</b>	22,8	15,2	15,6
<b>Rede 10</b>	11,2	3,6	12,8
<b>Média</b>	<b>12,28</b>	<b>12,0</b>	<b>18,4</b>

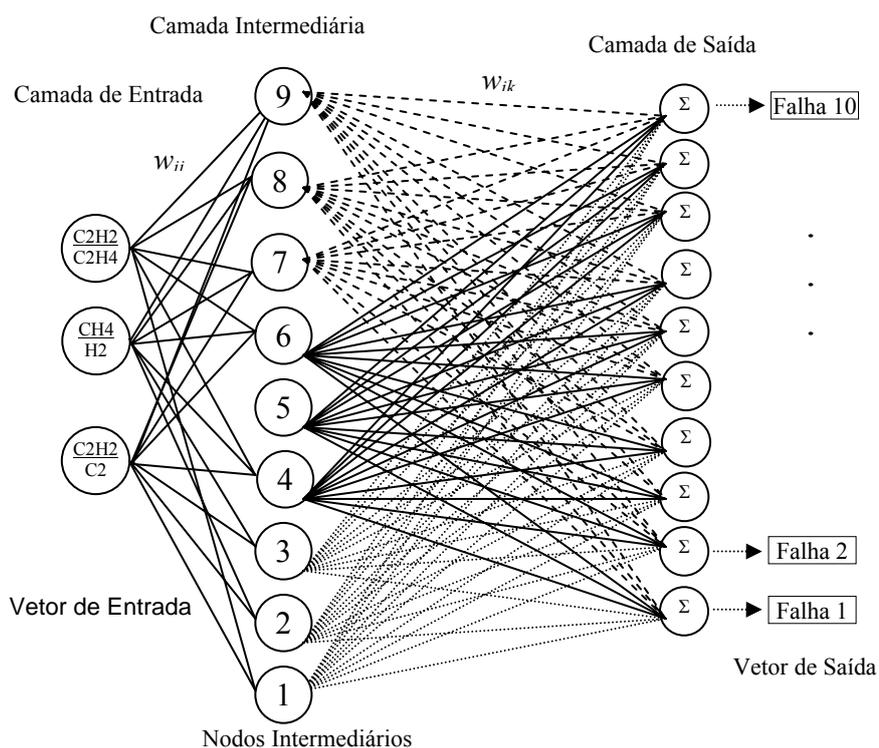
Na Tabela 14 é apresentada a análise do total de acerto por diagnóstico da rede escolhida. Nela, pode-se observar que o erro máximo de classificação por diagnóstico ocorreu na classe correspondente ao código IEC 001, sendo de 8%. O percentual de acerto total foi de 98% tendo a rede acertado 245 dos 250 diagnósticos que foram propostos, esse valor encontra-se na faixa de artigos que abordaram o mesmo problema utilizando redes neurais.

**Tabela 14** – Percentual de acerto na classificação.

<b>Diagnóstico</b>	<b>Código IEC</b>	<b>Acerto</b>	
		<b>Nº de padrões</b>	<b>Percentual (%)</b>
Sem Falha	000	25	100
Defeito térmico de baixa temperatura	001	23	92

(<150° C)			
Descargas Parciais de Pequena Densidade de Energia	010	24	96
Defeito térmico de baixa temperatura (150-300° C)	020	25	100
Defeito térmico de média temperatura (300-700° C)	021	25	100
Defeito térmico de alta temperatura (>700° C)	022	24	96
Descargas de energia reduzida	101	25	100
Descargas de alta energia	102	25	100
Descargas parciais de alta densidade de energia	110	25	100
Descargas de energia reduzida	202	24	96
Total	-	245	98

Na Figura 10 é possível observar uma ilustração da arquitetura final da rede escolhida com 3, 9 e 10 nodos na camada de entrada, intermediária e de saída, respectivamente.



**Figura 10-** Topologia da rede MLP selecionada

Nota-se que os nodos de entrada se conectam a todos os nodos da camada intermediária através de pesos e que estes, por sua vez, estão conectados a todos os nodos da camada de saída onde o ajuste dos pesos durante o treinamento da rede neural irá contribuir para que a saída seja incentivada ou penalizada a depender do sinal de entrada.

Durante o treinamento os pesos das conexões são ajustados de forma que sejam privilegiados ou penalizados para fornecer a saída compatível com o valor do vetor de entrada.

## 6. Implementação Lógica Fuzzy

A lógica fuzzy- LF -foi escolhida para implementar uma ferramenta de diagnóstico dos transformadores através da análise cromatográfica do óleo isolante devido a sua rápida adaptação a problemas que envolvem o conhecimento e a experiência adquiridos pelos seres humanos. Devido à falta de necessidade de treinamentos posteriores e graças à quantidade não suficiente de dados para um treinamento efetivo das redes neurais, a lógica fuzzy pôde ser adotada como alternativa à utilização de redes neurais.

Um das grandes vantagens do desenvolvimento da ferramenta com a utilização de lógica fuzzy é que os dados de entrada não precisam mais de pré-processamento, além disso, a ferramenta pode ser desenvolvida para diagnosticar duas falhas incipientes ocorrendo simultaneamente no transformador.

A utilização de lógica fuzzy para resolução desse tipo de problema vem sendo estudado por outros pesquisadores ([24]e[25]), tendo se mostrado bastante promissora. Essa técnica de inteligência artificial avalia o grau de pertinência de determinado elemento pertencer a determinadas funções membro para estabelecer uma probabilidade de o elemento fazer parte de um determinado grupo.

Essa característica foi utilizada para estimar quais concentrações dos gases pertencem a determinados diagnósticos propostos pelos métodos apresentados. De maneira que foi possível desenvolver uma ferramenta capaz de diagnosticar o transformador, com base nos métodos, a partir da concentração dos gases obtida durante a análise cromatográfica do óleo isolante.

Com a ferramenta desenvolvida foi possível avaliar os equipamentos com base nos métodos propostos e também aproximar para algum diagnóstico conhecido os casos em que os métodos diagnosticam como não tabelados, ou seja, os casos que não possuem uma concentração dos gases não contemplada no desenvolvimento do método. A criação da ferramenta para aproximar diagnósticos para as amostras ditas não tabeladas foi baseada, principalmente, nos laudos emitidos pelas empresas contratadas para analisar o óleo isolante.

Os dados utilizados durante o desenvolvimento do Sistema de Inferência Fuzzy – SIF – fazem parte do mesmo banco de dados de onde foram extraídos os dados para treinamento, validação e teste das redes neurais, de modo que é possível uma comparação direta entre as ferramentas desenvolvidas.

## 6.1. Desenvolvimento do SIF

Em linguagem computacional foi desenvolvido no Matlab® um script de implementação da lógica fuzzy, no qual é possível, a partir das relações dos gases contidos na análise cromatográfica do óleo isolante, fornecer um diagnóstico compatível com os métodos IEC e ROGERS. Em caso de limites de concentração de gases não tabelados pelos métodos, o SIF é capaz ainda de aproximar a análise sugerindo um diagnóstico mais provável, o qual indica a tendência de evolução do defeito incipiente deste transformador.

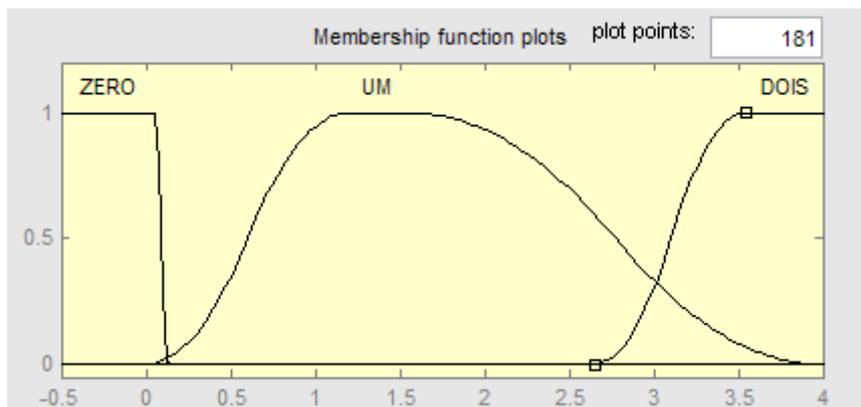
### 6.1.1. Método IEC

O limite estabelecido para cada relação gasosa ( $C_2H_2/C_2H_4$ ,  $CH_4/H_2$  ou  $C_2H_4/C_2H_6$ ), foi substituído por funções membros (ZERO, UM E DOIS) onde cada função representa uma faixa de intervalo de valores, inferior e/ou superior, que gera o código que remete a um dos diagnósticos mostrado na Tabela 3.

Portanto, a fuzzificação da relação dos gases nestas funções membros atribui valor de pertinência a cada valor de cada relação gasosa, tornando-as mais próximas possível do tabelamento pelo método IEC. Desse modo foram obtidas as três funções membros para cada relação gasosa, parametrizadas de forma heurística com o objetivo de satisfazer os limites estabelecidos na norma.

Com base nas faixas de valores impostas pelo método foram também definidas as formas das funções, por exemplo, a relação  $R_3=C_2H_4/C_2H_6$  não possui um limite inferior quando  $R_3 < 1$ , logo foi criada a função membro “Zero” para  $R_3$  sendo do tipo Zmf que tem por característica ser constante abaixo de determinado valor, sendo um tipo que se aproxima bastante do formato da limitação e tendo apenas dois limites superiores.

Essas funções escolhidas para substituir as faixas de limites propostos pelo método e devem indicar a probabilidade de um determinado valor da relação de concentração dos gases pertencer ao código que a função descreve, de forma que cada relação gasosa não apenas gerará ou não um determinado código, mas será representado por um percentual em todos os códigos dependendo da função membro.



**Figura 11** – Funções membro para a relação R1

A Figura 11 mostra as funções membro encontradas para a codificação da relação dos gases  $C_2H_2/C_2H_4$ , onde o código Zero é representado por uma função do tipo Zmf, o código um pelo tipo Pimf e o código Dois por uma função tipo Smf.

**Tabela 15** - Codificação IEC

$R_1 < 0,1$	0
$0,1 \leq R_1 \leq 3$	1
$R_1 > 3$	2

O exemplo abaixo ilustra a codificação gerada através da utilização das funções membro. Um valor foi escolhido para melhor visualização de como são gerados os valores percentuais da amostra a partir das funções membro. Escolhendo  $R_1=2,9$  e analisando a Tabela 15 é possível visualizar que o código correspondente gerado seria 1, porém na nova codificação através da lógica fuzzy e utilizando as funções membro a codificação deixa de ser única e passa a ter valor entre zero e um para cada uma das funções membro que representa os códigos, como pode ser visto na Figura 12, onde  $R_{1\_Zero}=0$ ,  $R_{1Um}= 0,4$  e  $R_{1\_Dois}= 0,15$ .

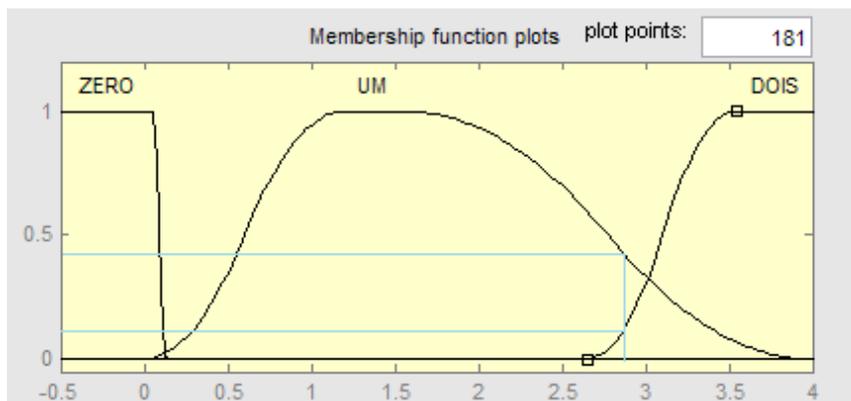


Figura 12 – Representação da metodologia para  $R_1=2,9$ .

A partir da composição dos valores de todas as relações gasosas,  $R_1=C_2H_2/C_2H_4$ ,  $R_2=CH_4/H_2$  e  $R_3=C_2H_4/C_2H_6$  em todas as funções membros (Zero, Um e Dois), são estabelecidas as probabilidades que indicam o grau de pertinência do resultado avaliado pertencer a determinado diagnóstico de acordo com os códigos mostrados na Tabela 4 e utilizando as seguintes funções:

$$F(0) = R_{1,Zero} + R_{2,zero} + R_{3,zero} \quad (42)$$

$$F(1) = R_{1,Zero} + R_{2,zero} + R_{3,um} \quad (43)$$

$$F(2) = R_{1,Zero} + R_{2,um} + R_{3,zero} \quad (44)$$

$$F(3) = R_{1,zero} + R_{2,dois} + R_{3,zero} \quad (45)$$

$$F(4) = R_{1,zero} + R_{2,dois} + R_{3,um} \quad (46)$$

$$F(5) = R_{1\_Zero} + R_{2\_Dois} + R_{3\_Dois} \quad (47)$$

$$F(6) = \text{máximo} ((R_{1\_Um} + R_{2\_Zero} + R_{3\_Um}) \text{ ou } (R_{1\_Um} + R_{2\_Zero} + R_{3\_Dois}) \text{ ou} \quad (48)$$

$$(R_{1\_Dois} + R_{2\_Zero} + R_{3\_Um}) \text{ ou } (R_{1\_Dois} + R_{2\_Zero} + R_{3\_Dois}))$$

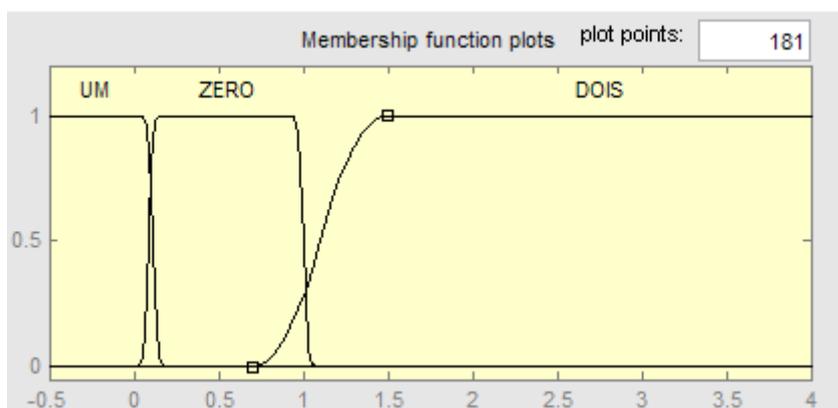
$$F(7) = R_{1\_Um} + R_{2\_Zero} + R_{3\_Dois} \quad (49)$$

$$F(8) = R_{1\_Um} + R_{2\_Um} + R_{3\_Zero} \quad (50)$$

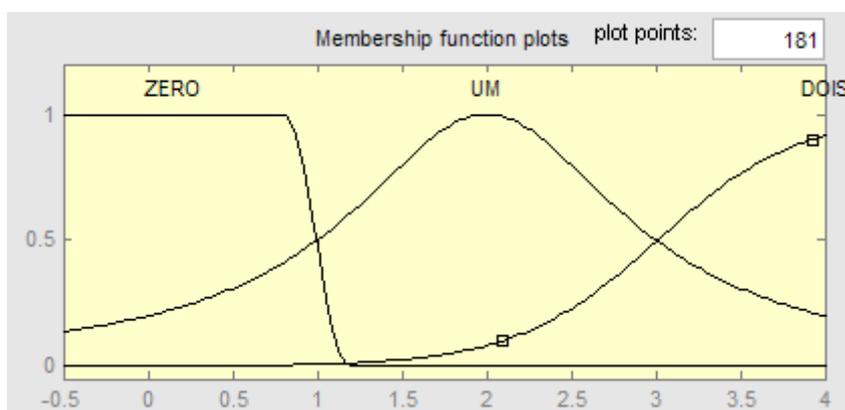
$F(0) = R_{1,Zero} + R_{2,zero} + R_{3,zero}$  onde,  $R_1$ ,  $R_2$  e  $R_3$  indicam as razões dos gases  $C_2H_2/C_2H_4$ ,  $CH_4/H_2$  e  $C_2H_4/C_2H_6$ , respectivamente.  $R_{n,MF}$  indica a fuzzificação da relação  $R_n$  na função membro MF que pode ser *Zero*, *Um* ou *Dois*, isto é,  $R_{1,Zero}$  indica a fuzzificação da relação  $R_1$  na função membro *Zero*. A função  $F(6)$  indica o diagnóstico Arco ou Descargas de baixa energia que pode ser diagnosticado através de dois códigos, logo foi utilizada a maior das probabilidades de ocorrência entre os dois códigos para indicar essa falha incipiente.

Com os valores de cada uma das funções mostradas acima as amostras foram diagnosticadas como aquela função que apresentou o maior valor. Nos casos em que foram apresentados valores iguais para mais de um defeito o diagnóstico sugerido pela metodologia implementada foi a possível ocorrência de duas falhas incipientes simultâneas no transformador. Os casos diagnosticados dessa maneira foram sempre casos onde o método não possuía tabelamento e a existência de defeitos simultâneos pode explicar a presença de concentrações gasosas atípicas.

A Figura 13 e a Figura 14 exibem, respectivamente, as funções membros criadas para as relações dos gases  $R_2$  e  $R_3$ , onde para a relação  $R_2$  a função “Zero” é do tipo Dsigmf, a função “Um” é Zmf e a função “Dois” do tipo Smf. Já para a relação  $R_3$  foram utilizados os tipos Zmf, Gbellmf e Sigmf para representar as funções “Zero”, “Um” e “Dois”, respectivamente.



**Figura 13** - Funções membro para a relação  $R_2$



**Figura 14** - Funções membro para a relação  $R_3$

### 6.1.2. Método ROGERS

A implementação criada para método ROGERS usou a mesma metodologia já utilizada para implementar o método IEC, sendo que utilizando as relações gasosas, as faixas de valores e a codificação sugerida no novo método. A escolha dos tipos de função membro utilizada também se baseou nas características dos limites fornecidos pelo método.

As relações utilizadas por esses métodos são  $R_1=CH_4/H_2$ ,  $R_2=C_2H_6/CH_4$ ,  $R_3=C_2H_4/C_2H_6$  e  $R_4=C_2H_2/C_2H_4$ , porém diferente do método IEC as faixas de valores tem quantidades diferentes de intervalos para cada relação dos gases, como pode ser visto na Tabela 6. As funções membro foram criadas baseando-se nos códigos que são gerados a partir da distribuição na faixa de valores, podendo a relação ser representada por quatro, três ou duas funções membro.

A Figura 15 exhibe a codificação da relação  $CH_4/H_2$  que possui quatro faixas de valores estabelecidos, onde os tipos de função membro utilizados foram Zmf, Dsigmf, Pimf e Sigmf para os códigos Cinco, Zero, Um e Dois respectivamente.

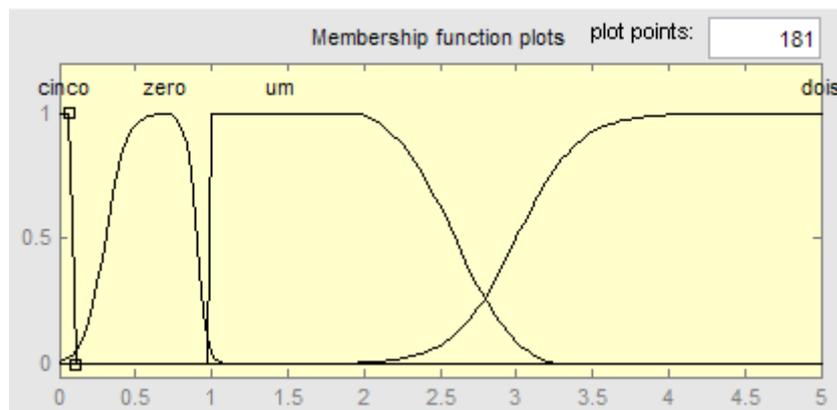


Figura 15– Funções membro para R1 ROGERS.

Da mesma forma utilizada no método IEC, foram criadas funções que representam o percentual de determinado valor da relação do gás pertencer às funções membro criadas. Para o método ROGERS foram utilizadas as seguintes funções:

$$F(0) = R_{1\_Zero} + R_{2\_Zero} + R_{3\_Zero} + R_{4\_Zero} \quad (51)$$

$$F(1) = R_{1\_Zero} + R_{2\_Zero} + R_{3\_Zero} + R_{4\_Um} \quad (52)$$

$$F(2) = R_{1\_Zero} + R_{2\_Zero} + R_{3\_Um} + R_{4\_Zero} \quad (53)$$

$$F(3) = \text{máximo}((R_{1\_Zero} + R_{2\_Zero} + R_{3\_Um} + R_{4\_Um}) \text{ ou } (R_{1\_Zero} + R_{2\_Zero} + R_{3\_Um} + R_{4\_Dois}) \text{ ou } (R_{1\_Zero} + R_{2\_Zero} + R_{3\_Dois} + R_{4\_Um})) \quad (54)$$

$$F(4) = R_{1\_Zero} + R_{2\_Zero} + R_{3\_Dois} + R_{4\_Dois} \quad (55)$$

$$F(5) = \text{máximo}((R_{1\_Zero} + R_{2\_Um} + R_{3\_Zero} + R_{4\_Zero}) \text{ ou } (R_{1\_Um} + R_{2\_Um} + R_{3\_Zero} + R_{4\_Zero}) \text{ ou } (R_{1\_Dois} + R_{2\_Um} + R_{3\_Zero} + R_{4\_Zero})); \quad (56)$$

$$F(6) = \text{máximo}((R_{1\_Um} + R_{2\_Zero} + R_{3\_Zero} + R_{4\_Zero}) \text{ ou } (R_{1\_Dois} + R_{2\_Zero} + R_{3\_Zero} + R_{4\_Zero})); \quad (57)$$

$$F(7) = R_{1\_Um} + R_{2\_Zero} + R_{3\_Um} + R_{4\_Zero} \quad (58)$$

$$F(8) = R_{1\_Um} + R_{2\_Zero} + R_{3\_Dois} + R_{4\_Zero}; \quad (59)$$

$$F(9) = R_{1\_Cinco} + R_{2\_Zero} + R_{3\_Zero} + R_{4\_Zero}; \quad (60)$$

$$F(10) = \text{máximo} (R_{1\_Cinco} + R_{2\_Zero} + R_{3\_Zero} + R_{4\_Um}) \text{ ou } (R_{1\_Cinco} + R_{2\_Zero} + R_{3\_Zero} + R_{4\_Dois}); \quad (61)$$

Onde os diagnósticos que podem ser encontrados pela formação de dois ou mais códigos foram diagnosticados de forma que a maior probabilidade dos códigos possíveis foi utilizada como a probabilidade do defeito incipiente. A partir dos valores resultantes dessas funções a amostra é diagnosticada como aquela que possuir o maior valor e em casos não tabelados a amostra pode ser diagnosticada com duas faltas incipientes.

As Figura 16, 17 e 18 abaixo trazem as funções membro que representam as relações  $R_2$ ,  $R_3$  e  $R_4$ , respectivamente.

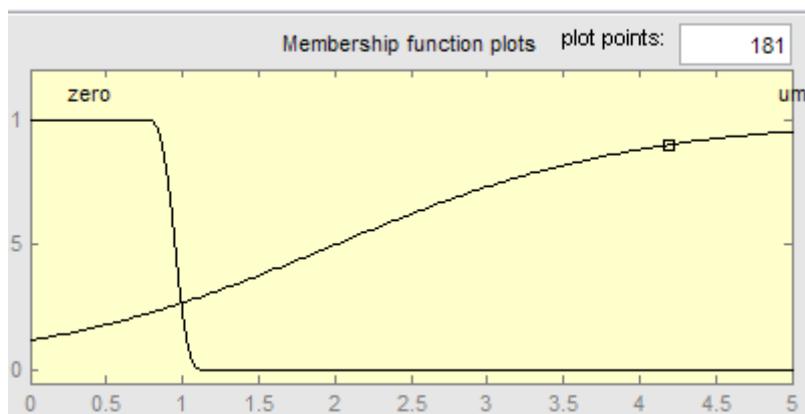
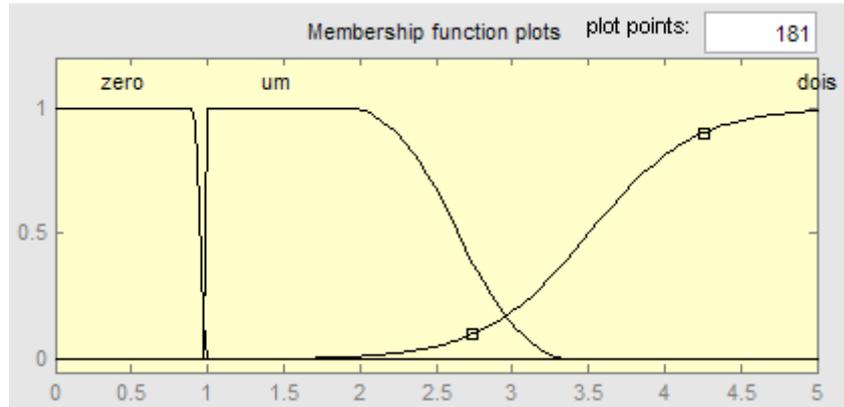
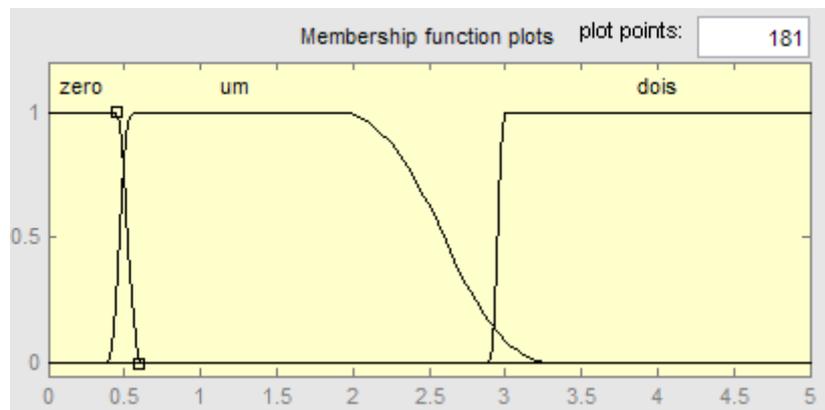


Figura 16 – Funções membro para R2 ROGERS.



**Figura 17** – Funções membro para R3 ROGERS.



**Figura 18** – Funções membro para R4 ROGERS.

## 7. Resultados

Apesar de promissores os resultados obtidos pelas simulações das redes neurais foram utilizados somente a título de comparação, já que a implementação utilizando lógica Fuzzy trouxe resultados mais contundentes com os diagnósticos fornecidos pelos métodos, como poderá ser observado ao longo desse capítulo.

Pela falta de necessidade de treinamento posterior, pois os métodos são fixos, e pela simplicidade do método adotado a utilização da Lógica Fuzzy trouxe vantagens significativas na implementação dos métodos de diagnósticos, eliminando o pré-processamento dos dados e a simulação dos dados tornou-se mais simples.

Em linguagem computacional foi desenvolvido um software, no qual é possível a partir das relações dos devidos gases contidos na análise cromatográfica do óleo isolante de um transformador de potência, fornecer um diagnóstico compatível com os dos métodos IEC e/ou ROGERS. Possui ainda a funcionalidade de aproximar alguns diagnósticos, cujos limites de concentração dos gases não possuem tabelamento através das normas, para um ou mais diagnósticos tabelados já que podem existir falhas simultâneas ocorrendo no transformador.

### 7.1. Lógica Fuzzy versus Rede Neural

Neste tópico serão apresentados os resultados obtidos com a simulação das redes neurais criadas e discutidas no Capítulo 0 e compará-los com os resultados encontrados com o Sistema de Inferência Fuzzy.

**Tabela 16** – Percentual de acerto na classificação por diagnóstico

DIAGNÓSTICO	CÓDIGO IEC	Erros		
		Nº total de padrões	No. padrões	Percentual (%)
Sem Falha	000	571	63	11,03
Defeito térmico de baixa temperatura <150° C	001	436	30	6,88
Descargas Parciais de Pequena Densidade de Energia	010	237	0	0
Defeito térmico de baixa	020	615	40	6,5

temperatura 150-300° C				
Defeito térmico de média temperatura 300-700° C	021	208	1	0,48
Defeito térmico de alta temperatura >700° C	022	232	9	3,88
Descargas de energia reduzida	101	41	2	4,88
Descargas de alta energia	102	143	7	4,9
Descargas parciais de alta densidade de energia	110	9	0	0
Descargas de energia reduzida	202	99	3	3,03
<b>Total</b>	-	2591	155	5,99

O total de erro por diagnóstico da rede selecionada é apresentada na Tabela 16. Nela, pode-se observar que o erro máximo de classificação por diagnóstico ocorreu na classe correspondente ao código IEC 000 (Sem falha), sendo de 11,03%. O percentual total de erro foi de 5,99%, cujo valor encontra-se dentro da faixa de erro dos resultados apresentados por artigos que abordaram o mesmo problema utilizando redes neurais ([22], [27]e[28]).

A partir dos resultados da lógica fuzzy foi possível classificar todos os diagnósticos pelas normas IEC e ROGERS com 100% de acerto e ainda aproximar os casos sem tabelamento para diagnósticos semelhantes.

A Tabela 17 expressa a quantidade de erros dos dados obtidos aplicando o SIF para a codificação IEC, onde as quantidades totais de dados analisados para cada tipo de falha e a quantidade de laudos mal diagnosticados pelo sistema são fornecidas nas colunas 2 e 3, respectivamente.

**Tabela 17– Resultados IEC**

<b>Codificação IEC</b>	<b>Quantidade de amostras</b>	<b>Erro</b>	<b>Erro (%)</b>
000	573	0	0
001	436	0	0
010	237	0	0
020	615	0	0
021	208	0	0

022	232	0	0
102	143	0	0
110	9	0	0
101			
202	140	0	0

Observa-se na Tabela 17 que o SIF diagnosticou todos os casos corretamente, os quais foram confirmados pela norma da IEC.

A Tabela 18 mostra o resultado obtido pelo SIF para a codificação proposta pela IEEE (ROGERS).

**Tabela 18 – Resultados ROGERS**

<b>Codificação ROGERS</b>	<b>Quantidade de amostras</b>	<b>Erro</b>	<b>Erro (%)</b>
0000	200	0	0,00
0001	13	0	0,00
0010	276	0	0,00
0011			
0012	89	0	0,00
0021			
0022	89	0	0,00
1000			
2000	285	0	0,00
1010	131	0	0,00
1020	138	0	0,00
1100			
2100	723	0	0,00
0100			
5000	98	0	0,00
5001			
5002	15	0	0,00

## 7.2.Exemplos de resultados

Os resultados mostram que tanto a utilização das técnicas de IA com Rede Neural quanto com a Lógica Fuzzy para a identificação de diagnósticos de transformadores foi bem sucedida. Ambas conseguiram apresentar os mesmos defeitos indicados pela norma IEC, sendo que a Rede Neural demonstra uma grande percentagem de erro quando comparada com a Lógica Fuzzy.

A lógica Fuzzy além de apresentar um melhor desempenho, também permitiu o estudo de caso em vários transformadores, através da análise de históricos associados aos laudos, apresentando-se mais completa e confiável. Então, devido ao elevado desempenho do sistema de inferência Fuzzy em relação a qualquer uma das redes neurais desenvolvidas, foi tomada a decisão de implementar no sistema a técnica de IA que utiliza a inferência Fuzzy.

A Tabela 19 exemplifica o acompanhamento de um transformador, comparando o diagnóstico sugerido pelo método IEC com o indicado pela SIF.

**Tabela 19– Acompanhamento de diagnóstico do Transformador**

	<b>Data da amostragem</b>	<b>Diagnóstico IEC</b>	<b>Diagnóstico SIF</b>
Transformador: 02T1 Ano: 1989 Série: 139413 Potência: 10000 kVA Tensão: 69kV Fabricante: Eletroputere Subestação: Afogados da Ingazeira	21/5/2003	Não Tabela	Defeito térmico de alta temperatura
	16/3/2004	Não Tabela	Defeito térmico de alta temperatura
	23/6/2004	Não Tabela	Defeito térmico de alta temperatura
	28/9/2004	Não Tabela	Defeito térmico de baixa temperatura
	9/12/2004	Não Tabela	Defeito térmico de alta temperatura
	3/3/2005	Não Tabela	Defeito térmico de alta temperatura
	14/6/2005	Não Tabela	Defeito térmico de alta temperatura
	12/9/2005	Não Tabela	Defeito térmico de alta temperatura
	27/12/2005	Não Tabela	Defeito térmico de alta temperatura
	17/3/2006	Não Tabela	Defeito térmico de alta temperatura
	27/9/2006	Defeito térmico de alta temperatura	Defeito térmico de alta temperatura
12/6/2007	Defeito térmico de alta temperatura	Defeito térmico de alta temperatura	

Nota-se que ao longo dos anos as amostras extraídas e analisadas através do método IEC tiveram, em sua grande maioria, um não tabelamento das falhas incipientes que podiam está ocorrendo no transformador. Porém como o SIF sempre aproxima um diagnóstico para a amostra, ele vinha sempre diagnosticando o transformador com um defeito térmico de alta temperatura, após algumas novas amostras, correspondendo a anos de operação, o método diagnosticou o equipamento com o mesmo defeito que já vinha sendo sugerido pelo SIF.

Isso foi observado em diversas análises em diferentes transformadores e pode ser útil para a equipe responsável pelo planejamento e execução das manutenções em transformadores de potência, ajudando na tomada de decisão da equipe e evitando a ocorrência de acidentes e da possível inutilização do equipamento.

**Tabela 20** – Exemplo de aplicação da metodologia

	<b>Data da amostragem</b>	<b>Diagnóstico IEC</b>	<b>Diagnóstico SIF</b>
<b>Transformador:</b> 02T1 <b>Ano:</b> 2001 <b>Série:</b> 184662 <b>Potência:</b> 20000 kVA <b>Tensão:</b> 69kV <b>Fabricante:</b> WEG <b>Subestação:</b> João de Barros	30/11/2001	Não Tabelado	Descargas de energia reduzida
	27/12/2001	Não Tabelado	Descargas de energia reduzida
	27/12/2001	Não Tabelado	Descargas de energia reduzida
	28/1/2002	Não Tabelado	Descargas de energia reduzida
	2/12/2002	Não Tabelado	Descargas de energia reduzida
	25/3/2003	Descargas de energia reduzida	Descargas de energia reduzida
	10/3/2004	Sem falha	Sem falha
	11/3/2005	Defeito térmico de baixa temperatura	Defeito térmico de baixa temperatura
	14/3/2005	Não Tabelado	Descargas de energia reduzida
	14/3/2005	Não Tabelado	Descargas de energia reduzida
	10/10/2007	Descargas de energia reduzida	Descargas de energia reduzida

O exemplo acima ilustra um caso em que o equipamento vinha sendo avaliado em um intervalo de tempo muito menor do que normalmente utilizado para analisar o óleo isolante, em vista da sua recente reenergização, por ter passado por uma circulação da carga total de óleo no tanque usando termo vácuo.

Nos diagnósticos mostrados na Tabela 20 observa-se que após pouco tempo de energização o método da IEC não conseguia nenhum apontamento de falhas incipientes ocorrendo no transformador, pois a geração dos gases no momento da energização é normalmente atípica, porém o SIF apontava para a existência de descargas de energia reduzida. Diagnóstico esse que veio ser tabelado pelo método IEC após algumas novas análises, após a confirmação desse diagnóstico o óleo sofreu um tratamento de

desgaseificação e em uma coleta após esse tratamento teve um diagnóstico de sem falha ocorrendo em seu interior, porém após novas amostragens os defeitos foram novamente evoluindo e passando de defeito térmico de baixa temperatura para descargas de energia reduzida, o que mostra que algum sobreaquecimento interno no equipamento está provocando essa geração de gases.

## 8. Conclusão

A presente dissertação veio mostrar o desenvolvimento da ferramenta para diagnóstico de falhas incipientes em transformadores de potência, através da análise cromatográfica do óleo isolante e apresentar os resultados da pesquisa desenvolvida em parceria com a CELPE.

No capítulo 2 foi apresentada a base teórica envolvendo o histórico das redes neurais e o seu funcionamento. Foram apresentados os elementos básicos que compõem uma rede neural e o funcionamento de algum deles. O funcionamento dos neurônios artificiais e exemplos de suas funções de ativação podem ser vistos neste mesmo capítulo, que traz ainda uma descrição dos principais algoritmos de treinamento utilizados para o desenvolvimento de Redes Neurais Artificiais, incluindo o algoritmo Levenberg-Maquardt selecionado para ser utilizado no desenvolvimento dessa pesquisa.

O capítulo 3 trouxe a teoria dos conjuntos Fuzzy, neste capítulo é mostrada a origem e um breve histórico do desenvolvimento dessa ferramenta, também as propriedades e operações matemáticas que podem ser utilizadas nesse conjunto, além de mostrar as principais funções de pertinência que podem ser utilizadas e os sistemas de inferência mais comumente encontrados.

O capítulo 4 mostrou uma descrição simplificada da importância do óleo isolante para os transformadores e mostrou-se a origem da ferramenta de análise do estado do transformador baseado na análise cromatográfica do óleo isolante. Foram mostradas as principais técnicas utilizadas para a interpretação das concentrações gasosas fornecidas na análise cromatográfica, a saber, o método proposto pela IEC e o método de ROGERS que foi adotado na norma IEEE. Neste capítulo é explicada a forma como os métodos constroem seus diagnósticos e quais as concentrações dos gases são utilizadas para formar as relações gasosas necessárias ao diagnóstico.

Nos capítulos 5 e 6 foram exibidas as implementações necessárias para a construção da ferramenta desenvolvida, tanto para a rede neural como para a utilização da lógica fuzzy. Nestes capítulos são descritos os ajustes necessários para a utilização dos dados, como a forma de pré-processamento utilizada para a rede neural e a criação das funções membro para a lógica Fuzzy. Também são mostradas as formas como foram elaborados os dados de entrada e saída da rede neural e a separação dos dados para a montagem dos conjuntos de treinamento, validação e teste. A arquitetura da rede neural selecionada e os critérios utilizados para essa seleção também podem ser visto neste

capítulo. Na utilização da lógica Fuzzy é mostrado o desenvolvimento da ferramenta e as funções criadas para exibir a pertinência de uma amostragem pertencer a determinado diagnóstico, além de mostrar os tipos de funções membro utilizado para cada relação gasosa, tanto para o método IEC quanto para o método ROGERS.

O capítulo 7 traz os principais resultados obtidos com a simulação da ferramenta criada e uma comparação feita entre a simulação realizada com o resultado da rede neural criada e o da simulação da lógica fuzzy. Também foram citados exemplos das facilidades que puderam ser extraídas com a utilização dessa nova ferramenta.

### **8.1. Trabalhos futuros**

Como trabalhos futuros para melhorar o desempenho da ferramenta desenvolvida, obtendo um resultado mais compatível com o que está acontecendo no equipamento são propostos:.

- Utilizar novos métodos de diagnósticos como o método do gás chave e o triângulo de Duval para complementar os diagnósticos sugeridos pela ferramenta;
- Desenvolver a ferramenta para que ela possa unir, automaticamente, os resultados de diversos métodos em um resultado completo;
- Utilizar os dados coletados para desenvolver uma metodologia específica por família de transformadores;
- Avaliar a utilização da ferramenta em equipamentos de diagnóstico on-line dos transformadores de potência.

## Revisão Bibliográfica

- [1] BARROS FILHO, Vespúcio Nunes de Alencar. Dissertação de Mestrado **"Monitoramento e Diagnóstico de Estados de Grandes Transformadores com Ênfase para o Diagnóstico pela Cromatografia com o Uso de Redes Neurais"** - UFPE, Recife-PE, Setembro, 2003.
  
- [2] IEC - International Electrotechnical Commission - **"Interpretation the analysis of gases in transformer and other Oil-filled impregnated Electrical Equipment in Service"** , CEI - IEC - 599, First Edition, 1978.
  
- [3] IEC - International Electrotechnical Commission - **"Mineral oil-impregnated electrical equipment in service Guide to the interpretation of dissolved and free gases analysis"**, CEI - IEC - 60599, Second Edition, 1999.
  
- [4] IEEE – Institute of Electrical and Electronics Engineers – **“Guide for the detection and determination of generated gases in oil-immersed transformers and their relation to the serviceability of the equipment”**, ANSI/IEEE C57.104-1978.
  
- [5] IEEE – Institute of Electrical and Electronics Engineers – **“Guide for the interpretation of gases generated in oil-immersed transformers”**, IEEE std C57.104-1991.
  
- [6] Teixeira, R. de A.; Corrêa, M.V.; Assis, J. C. de Oliveira **Redes Neuro Fuzzy Algoritmo de Treinamento de Aplicações.**
  
- [7] FAUSETT, L.; 1994. **"Fundamentals of Neural Networks"**. New Jersey: Prentice Hall Inc.FAUSETT, L.; 1994.
  
- [8] [Wasserman P.D., **“Neural Computing: Theory and Practice”**, New York,

Van Nostrand Reinhold, 1993.]

- [9] BRAGA, Antônio; CARVALHO, André; LUDERMIR, Teresa. “**Redes Neurais Artificiais: Teoria e Aplicações**”. Rio de Janeiro: LTC, 2000.
- [10] M. Riedmiller and H. Braun, "A **Direct adaptive Method for faster Backpropagation Learning: The RPROP Algorithm**", IEEE International Conference on Neural Networks, vol.1, April 1993, pp.586-591.
- [11] Levenberg, K. "A **Method for the Solution of Certain Problems in Least Squares**." Quart. Appl. Math. 2, 164-168, 1944.
- [12] Marquardt, D. "An **Algorithm for Least-Squares Estimation of Nonlinear Parameters**." SIAM J. Appl. Math. 11, 431-441, 1963.
- [13] Gill, P. E. and Murray, W. "Algorithms for the solution of the **nonlinear least-squares problem**", SIAM J. Numer. Anal. 15 [5] 977-992, 1978.
- [14] BEZDEK, J. C.; PAL, S. K. “**Fuzzy Models for Pattern Recognition**”, IEEE Press, New York, 1992.
- [15] ZADEH, L. A. Fuzzy sets. **Information and Control**, v. 8, p. 338-353, 1965.
- [16] ANTUNES, J. F. G. **Aplicação de Lógica Fuzzy para estimativa de área plantada da cultura de soja utilizando imagens AVHRR-NOAA**. 2005. 105f. Dissertação (Mestrado em Engenharia agrícola) – Universidade Estadual de Campinas, São Paulo, 2005.
- [17] TANSCHKEIT, R. . Sistemas fuzzy. In: **VI Simpósio Brasileiro de Automação Inteligente**, 2003, Bauru, SP. Anais de Minicursos do VI SBAI, 2003. p. 35 páginas.

- [18] Marinho, M. H. da N.; Andrade Filho, M. G. de; Ballini, R.; Santos, E. F. dos; Soares Filho, S. **“Previsão de Vazões Afluentes Utilizando Modelos Periódicos Auto-Regressivos, Redes Neurais e Redes Neurais Nebulosas”**.
- [19] **Sistema de Monitoração Geração de Transformadores**, ABB. Disponível em: [http://www.eyeproject.cl/files/42dc39\\_ABB-BR-34-ABBTMonitor-SistMonitoracao.pdf](http://www.eyeproject.cl/files/42dc39_ABB-BR-34-ABBTMonitor-SistMonitoracao.pdf)>. Acesso em: 29 jul. 2011.
- [20] Rezende, Solange O. **“Sistemas inteligentes: fundamentos e aplicações”**, Barueri, ISBN 85-204-1683-7, Manole ,2005.
- [21] PRECHELT, L. Proben1 – **“A set of Neural Network Benchmark Problems and Benchmarking Rules”**. Technical Report21/94, Fakultät für Informatik, Universität Karlsruhe, Germany, 1994.
- [22] [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, October, 2001.]
- [23] HAGAN, M.T. & MENHAJ, **“M. Training Feedforward Networks with the Marquardt Algorithm”**. IEEE Transactions on Neural Networks, vol. 5, no. 6, 1994
- [24] C. E. Lin, J. M. Ling, and C. L. Huang, **“An Expert System For Transformer Faults Diagnosis And Maintenance Using Dissolved Gas Analysis”**, IEEE Trans. On Power Delivery, vol. 8, no.1, pp. 231-238 , January 1993.
- [25] Q. Su, C. Mi., L. L. Lai and P. Austin, **“A Fuzzy Dissolved Gas Analysis Method for Diagnosis of Multiple Incipient Faults in a Transformer”**, IEEE

Trans. On Power Systems, vol. 15, no.2, pp. 593-598, May 2000]

- [26] MATLAB, **The Language of Technical Computing**. Version 7.3 R2006b.
  
- [27] BARROS FILHO, Vespúcio Nunes de Alencar; AQUINO, Ronaldo Ribeiro Barbosa de; FERREIRA, Aida Araújo; SOUZA, Leonardo Tadeu Araújo. **“Monitoramento e Diagnóstico de Estado de Grandes Transformadores, com Ênfase para o Diagnóstico Pela Cromatografia com o Uso de Redes Neurais”**. In: IX Symposium of Specialists in Electric Operational and Expansion Planning, 2004, Rio de Janeiro. IX SEPOPE -SP-167. 2004. v. CD-ROM, p. 1-6.
  
- [28] BARROS FILHO, Vespúcio Nunes de Alencar; AQUINO, Ronaldo Ribeiro Barbosa de; FERREIRA, Aida Araújo. **“Diagnóstico de Estado de Transformadores através da Cromatografia e com o Uso de Redes Neurais”**. In: XVIII SNPTEE - Seminário Nacional de Produção e Transmissão de Energia Elétrica, 2005, Curitiba. XVIII SNPTEE - CD ROM. 2005. v. 1, p. 1-6.
  
- [29] Haykin,S. **“Redes Neurais – Princípios e práticas”**. Bookman, Segunda Edição, Porto Alegre –RS- 2001.