



Pontifícia Universidade Católica de Minas Gerais
Programa de Pós Graduação em Engenharia Elétrica

**SISTEMAS DE DIAGNOSTICO DE FALTAS EM
TRANSFORMADORES DE POTÊNCIA UTILIZANDO ANÁLISE
DE GASES DISSOLVIDOS E TÉCNICAS DE INTELIGÊNCIA
COMPUTACIONAL**

MICHEL BORTOLINI HELL

DISSERTAÇÃO DE MESTRADO PPGEE-XX / 2002
ORIENTADOR : PROF. DR. PYRAMO PIRES DA COSTA JÚNIOR

NOVEMBRO / 2002



Pontifícia Universidade Católica de Minas Gerais
Programa de Pós Graduação em Engenharia Elétrica

MICHEL BORTOLINI HELL

**SISTEMAS DE DIAGNOSTICO DE FALTAS EM
TRANSFORMADORES DE POTÊNCIA UTILIZANDO
ANÁLISE DE GASES DISSOLVIDOS E TÉCNICAS
DE INTELIGÊNCIA COMPUTACIONAL**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, da Pontifícia Universidade Católica de Minas Gerais, como requisito parcial para obtenção do grau de Mestre em Engenharia Elétrica, elaborada sob a orientação do Prof. Dr. Pyramo Pires da Costa Júnior.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
BELO HORIZONTE
2002

À Rafaela, amor da minha vida.

AINDA QUE EU FALASSE A LÍNGUA DOS HOMENS,
E FALASSE A LÍNGUA DOS ANJOS,
SEM AMOR EU NADA SERIA.

"1 Coríntios 13"

Agradecimentos

Ao orientador, professor Pyramo Pires Costa Júnior: pelo apoio, motivação e competência com que desempenha o seu trabalho, e enfim, pela grande participação no desenvolvimento desta dissertação.

Aos Professores do Programa de Pós-Graduação em Engenharia Elétrica da PUC-MG, em especial aos professores Petr e Reinaldo que, com seus conhecimentos, em muito contribuíram na realização deste trabalho.

Ao Marcos Flávio, pelo apoio e discussões no desenvolvimento desta dissertação.

À Companhia Energética de Minas Gerais (CEMIG), pela concessão do banco de dados e pelo financiamento da pesquisa

Ao Neymard, pela ajuda e esclarecimentos das informações fornecidas pela CEMIG.

Ao Álvaro e ao Ronald, pelo precioso apoio e por estarem sempre a disposição para superar as dificuldades impostas.

Ao Conselho Nacional de Pesquisa e Desenvolvimento (CNPq) pelo fomento.

Ao pessoal administrativo, pela atenção e presteza no atendimento.

Aos colegas do curso Sérgio, Daniel e aos alunos da iniciação científica, Guilherme, Luiz,

Frederico, José Eduardo e Mário, pela ajuda no desenvolvimento do protótipo de sistema de diagnóstico.

Aos meus pais, Amilton e Lizete, por suportarem a saudade em função do meu ideal.

Aos meus irmão, Charles e Milena, pelo incentivo.

Ao amigo Alessandro, pelos momentos de descontração que em muito ajudaram.

E em especial à Rafaela, a companheira de todas as horas, que me apoiou desde o começo e soube como ninguém suportar as ausências. Sem ela este trabalho não seria possível.

Conteúdo

Resumo	12
Notação e Definições	13
1 Considerações Iniciais	15
1.1 Introdução	15
1.2 Relevância	16
1.3 Escopo	17
1.4 Objetivos	17
1.5 Metodologia	18
1.6 Estrutura do Trabalho	19
2 Revisão de Literatura	20
2.1 Diagnóstico de Falhas em Transformadores	20
2.1.1 Tipos de falhas e suas causas	21
2.1.2 Gases Normalmente Encontrados Dissolvidos no Óleo Isolante	22
2.1.3 Decomposição do Óleo Isolante e dos Materiais Celulósicos	23
2.2 Critérios de Análise de Gases Dissolvidos	24
2.2.1 Método dos Gases Chave	25
2.2.2 Método das Razões de Dörnenburg	26

2.2.3	Método das Razões de Rogers	28
2.3	Ferramentas Para Diagnóstico de Faltas a Partir dos Critérios DGA	30
2.3.1	Lógica Fuzzy	30
2.3.2	Redes Neurais Artificiais	33
3	Fundamentação Teórica	36
3.1	Teoria de Conjuntos Fuzzy	37
3.1.1	Conjuntos Fuzzy	38
3.1.2	Definições e Operações	38
3.1.3	Variáveis Lingüísticas	41
3.1.4	Proposições Fuzzy	42
3.2	Sistemas de Inferência Fuzzy	43
3.2.1	Modelo Fuzzy de Mamdani	44
3.2.2	Modelo Fuzzy de Takagi e Sugeno	47
3.2.3	Modelo Fuzzy de Tsukamoto	48
3.3	Sistema de Inferência Fuzzy Neuro-Adaptativo	49
3.3.1	Arquitetura ANFIS	50
3.4	Redes Neurais de Múltiplas Camadas	52
3.4.1	O Algoritmo <i>Backpropagation</i>	53
3.5	Redes Neurais de Função de Base Radial	57
3.6	Mapas Auto Organizáveis	59
3.6.1	O Modelo de Kohonen	60
4	Implementação e Análise dos Sistemas de Diagnóstico Propostos	63
4.1	Base de Dados Utilizada	64

4.2	Sistema de Simulação de Diagnóstico	64
4.3	Técnicas de Inteligência Computacional Aplicadas ao Diagnóstico de Faltas em Transformadores de Potência	66
4.4	Rede Neural MLP Aplicada no Diagnóstico de Faltas em Transformadores de Potência	66
4.5	Rede Neural de Função de Base Radial Aplicada no Diagnóstico de Faltas em Transformadores de Potência	76
4.6	Modelo de Kohonen Aplicado no Diagnóstico de Faltas em Transformadores de Potência	77
4.6.1	ANFIS Aplicado no Diagnóstico de Faltas em Transformadores de Potência	81
4.7	Análise Comparativa dos Resultados Apresentados pelos Sistemas de Diagnóstico de Faltas em Transformadores de Potência	84
5	Conclusões e Propostas de Trabalhos Futuros	86
5.1	Conclusões	87
5.2	Proposição para Trabalhos Futuros	87
5.2.1	Agrupamento da Base de Dados Utilizada	94
5.3	Agrupamento da Base de Dados Utilizando o Algoritmo de Adeli-Hung	96
	Bibliografia	90
	Abstract	93
	Apêndices	94

Lista de Figuras

2.1	Fluxograma de Aplicação do Método das Razões de Dörnenberg	27
2.2	Relação Metano / Hidrogênio	32
2.3	Relação Acetileno / Etileno	33
2.4	Relação Etileno / Etano	33
2.5	Sistema de Inferência Fuzzy para diagnóstico de faltas em transformadores baseado no método de Rogers	33
2.6	RNA para diagnóstico dos principais tipos de falta proposta por Zhang	34
2.7	RNA para diagnóstico de degradação da celulose proposta por Zhang	35
2.8	Sistemas de RNA's proposta por Wang	35
3.1	Operações com conjuntos fuzzy	40
3.2	Variável lingüística <i>Temperatura</i>	41
3.3	Diagrama de Blocos de um Sistema de Inferência Fuzzy	44
3.4	Sistemas de inferência fuzzy de Mamdani	45
3.5	Métodos de desfuzificação	47
3.6	Sistemas de inferência fuzzy de Takagi e Sugeno de Primeira Ordem	48
3.7	Sistemas de Inferência Fuzzy de Tsukamoto	49
3.8	(a) Modelo fuzzy Sugeno de primeira ordem de duas entradas com duas regras; (b) Estrutura ANFIS correspondente	50

3.9	Diagrama ilustrativo de uma Rede Neural com duas camadas intermediárias . . .	53
3.10	Fluxo de sinal ilustrando o efeito da constante de momento	56
3.11	Estrutura Básica de uma RBFN	57
3.12	Modelo de Campo Receptivo	58
3.13	Exemplo de Função de Base Radial	58
3.14	Diagrama Ilustrativo de Mapa de Kohonen	60
3.15	Diagrama Ilustrativo de Camada de Kohonen	61
3.16	Aproximação de Vetores na Rede de Kohonen	62
4.1	Fluxograma do Sistema de Simulação de Diagnóstico	65
4.2	Exemplo de Função de Ativação	67
4.3	Problema de Mínimos Locais	68
4.4	Primeira Estrutura de Rede Neural MLP Implementada	69
4.5	Função de Transferência Sigmoidal	70
4.6	Estrutura da Segunda Rede Neural MLP Proposta	72
4.7	Estrutura da Terceira Rede Neural MLP Proposta	74
4.8	Estrutura da Rede Neural RBFN Proposta	76
4.9	Estrutura da Camada de Kohonen Inicialmente Proposta	78
4.10	Estrutura da Rede Neural de Kohonen Proposta	79
4.11	Resultados Obtidos para o Sistema de Diagnóstico Baseado na Rede Neural de Kohonen	80
4.12	Sistema de Inferência Fuzzy Implementado (FIS)	82
4.13	Estrutura ANFIS Proposta	83
5.1	Esquema Básico de Modelo Térmico de Transformadores de Potência	88

5.2	Esquema Básico de Modelo Térmico de Transformadores de Potência	89
5.3	Esquema Básico de Modelo Térmico de Transformadores de Potência	89
5.4	Avaliação por Famílias	97

Lista de Tabelas

2.1	Gases Normalmente encontrados no óleo	23
2.2	Limites de Concentração	26
2.3	Faixas de Valores Característicos das Relações de Gases Dissolvidos no óleo	27
2.4	Faixas de Valores Característicos das Relações de Gases Livres (relé ou almofada)	27
2.5	Código para Análise de Gases Dissolvidos em Óleo Mineral	29
4.1	Correlação entre os Tipos de Falta e a Saída da Primeira Rede Neural MLP	70
4.2	Parâmetros Usados na Primeira Rede Neural MLP	71
4.3	Correlação entre os Tipos de Falta e a Saída da Rede Neural MLP	72
4.4	Parâmetros Usados na Rede Neural MLP	73
4.5	Resultados Obtidos para a Segunda Rede Neural MLP	73
4.6	Parâmetros Usados na Rede Neural MLP	75
4.7	Resultados Obtidos para a Segunda Rede Neural MLP	75
4.8	Resultados Obtidos para a Rede Neural RBFN	77
4.9	Resultados Numéricos Obtidos para o Sistema de Diagnóstico Baseado na Rede Neural de Kohonen	81
4.10	Resultados Obtidos para o Sistema de Diagnóstico ANFIS	83
4.11	Resumo dos Resultados Numéricos Obtidos	85

Resumo

Neste trabalho tratou-se do problema de diagnosticar faltas incipientes em transformadores através da análise dos gases que se encontram dissolvidos no óleo isolante. Na solução do problema foram aplicadas técnicas de inteligência computacional com a finalidade de desenvolver sistemas de diagnóstico que pudessem reconhecer os padrões de defeito apresentados pelos transformadores analisados. O primeiro sistema desenvolvido foi baseado em uma rede neural *Perceptron* de múltiplas camadas, treinada pelo algoritmo *backpropagation*. O segundo sistema desenvolvido foi baseado em uma abordagem de rede neural de função de base radial. A terceira técnica aplicada resultou no desenvolvimento de uma rede neural de Kohonen. O quarto sistema desenvolvido baseava-se em um sistema de inferência fuzzy e o quinto sistema implementou um sistema de inferência fuzzy neuro-adaptativo. A partir dos resultados, conclui-se que a rede neural de Kohonen é mais indicada para a classificação de defeitos, obtendo um grande percentual de acertos com baixo custo computacional. O trabalho também apresenta uma aplicação do algoritmo de Adeli-Hung no problema de agrupamento dos transformadores analisados, como forma de melhor tratar a base de dados usada no desenvolvimento dos sistemas, obtendo assim melhores resultados quando validados através de diagnósticos reais coletados em campo.

Notação e Definições

- \mathbb{R}^n - espaço dos números reais de dimensão n
- $\min(a, b) = a \wedge b$ - mínimo entre a e b
- $\max(a, b) = a \vee b$ - máximo entre a e b
- $\| \cdot \|$ - norma *euclidiana* no espaço \mathbb{R}^n
- \bar{A} conjunto complemento de A
- ω_{ji} peso de conexão sináptica que liga o neurônio j da camada antecedente ao neurônio i da camada conseqüente.
- Falta: Alteração das condições de um item, equipamento ou sistema operacional, de importância suficiente para que sua função normal não seja satisfatória. Uma falta não torna o equipamento indisponível, mas se não reparada ou se não corrigida levará o equipamento à uma falha e à conseqüente indisponibilidade (BRANCO FILHO, 1996).
- Falha: Perda da capacidade de um item para realizar sua função específica. É a diminuição total ou parcial da capacidade de uma peça, componente ou equipamento de desempenhar a sua função durante um período de tempo, onde o item deverá sofrer manutenção ou ser substituído. A falha leva o item ao estado de indisponibilidade (BRANCO FILHO, 1996).

- Diagnóstico: Identificação de causa provável de uma falha ou falta, com a ajuda de dados levantados, experiência e raciocínio (BRANCO FILHO, 1996).

Capítulo 1

Considerações Iniciais

1.1 Introdução

Atualmente um dos mais importantes objetivos das pesquisas em engenharia elétrica é a utilização da energia, de forma racional e confiável, nos mais diversos segmentos da sociedade. Isto requer o estudo de diversas formas para se alcançar a melhor performance possível do sistema elétrico e dos equipamentos e dispositivos a ele agregados. Dentre estes equipamentos, os transformadores assumem especial importância, uma vez que as condições operativas do sistema podem ser altamente influenciadas por variações no comportamento destes equipamentos.

Neste contexto, o constante monitoramento dos transformadores torna-se essencial para uma correta avaliação do seu estado operativo, tendo como finalidade aumentar a confiabilidade do sistema e reduzir custos de operação.

A ocorrência de faltas tais como superaquecimento, arco ou descargas parciais, podem causar interrupções no fornecimento de energia, resultando em altos custos. Estas faltas são desencadeadas por esforços elétricos, térmicos e mecânicos aos quais os transformadores são submetidos durante o funcionamento. Durante a ocorrência dessas faltas processos de decomposição química do material isolante são desencadeados, resultando na geração de gases. O tipo, a quantidade e a proporção desses gases dependem do material degradado, do fenômeno responsável pela degradação e dos níveis de energia envolvidos na ação. Desta forma é possível caracterizar a falta e a sua severidade através da análise da composição dos gases que se encontram dissolvidos no óleo isolante, bem como acompanhar tendências de evolução das faltas incipientes podendo-se, assim, programar ações de manutenção a fim de impedir desgaste

excessivo do equipamento evitando a perda inesperada do transformador, aumentando, com isso, a confiabilidade do sistema, além de reduzir os custos de manutenção.

Diversos critérios para o diagnóstico de faltas em transformadores a partir da análise dos gases dissolvidos no óleo tem sido desenvolvidos e vem sendo largamente utilizados. Esses critérios, chamados DGA ¹, envolvem vários métodos, tais como o Método dos Gases Chave (FALLOU, 1975), o Método das Razões de Dörnenburg (DÖRNENBURG e STRITTMATTER, 1974), (FALLOU et al., 1970) e o Método das Razões de Rogers (ROGERS, 1975),(ROGERS, 1978). Entretanto, uma vez que estes métodos foram desenvolvidos a partir de análises estatísticas de bases de dados européias, seus resultados, quando aplicados a equipamentos instalados em locais onde as condições climáticas diferem daquelas observadas na Europa, apresentam erros significativos. Com a finalidade de reduzir estes erros diversos sistemas de diagnóstico baseados nestes métodos vem sendo desenvolvidos, usando principalmente técnicas de Inteligência Computacional (IC), tais como os apresentados em (DUKARM, 1993), (ZANG et al, 1996), (WANG et. al., 2000), (HUANF et. al., 1997). Neste contexto o uso de IC tem se mostrado bastante promissor, uma vez que dispensa análises estatísticas complexas de cada planta de transformadores a ser monitorada.

1.2 Relevância

O transformador de potência é considerado um dos mais importantes componentes em um sistema elétrico de potência. Estes equipamentos são formados por complexas estruturas de enrolamentos que estão submetidos a diversos esforços durante seu funcionamento. Com o passar dos anos, os transformadores em operação se tornam mais suscetíveis à ocorrência de faltas, devido ao processo natural de envelhecimento e aos esforços aos quais eles são submetidos. Com isso, torna-se imprescindível o acompanhamento do estado operativo do mesmo, com a finalidade de aumentar a segurança e a confiabilidade do sistema elétrico.

Conforme citato anteriormente, a análise de gases dissolvidos é uma poderosa ferramenta para diagnóstico de faltas em transformadores. Muitos métodos de diagnóstico tem sido

¹ Do inglês *Dissolved Gas Analysis*

desenvolvidos para a interpretação da presença dos gases dissolvidos. Esses critérios possibilitam encontrar relações entre os gases e as condições de faltas em transformadores. Porém, muitos desses diagnósticos necessitam de peritos para interpretar os resultados corretamente. Nas últimas décadas tem havido uma grande preocupação em se desenvolver metodologias associadas a análise de gases e técnicas de inteligência computacional para a obtenção de diagnóstico de faltas incipientes de forma automatizada, com a finalidade de servir como ferramenta de auxílio ao perito, aumentando sua percepção. Dentre as técnicas de inteligência computacional que vem sendo aplicadas ao diagnóstico de faltas em transformadores de potência, podemos citar os sistemas de inferência fuzzy (DUKARM, 1993), (HUANF et. al., 1997) e as redes neurais artificiais (RNA's) (ZANG et al, 1996), (WANG et. al., 2000), que são usadas desde que as relações ocultas entre os tipos de faltas e os gases dissolvidos possam ser reconhecidos pela RNA por meio do processo de treinamento.

Aplicando estas técnicas no desenvolvimento de sistemas de diagnóstico de faltas é possível obter um poderoso sistema de apoio à manutenção, com a finalidade de evitar perdas inesperadas de transformadores, aumentando a confiabilidade do sistema elétrico, bem como reduzindo os custos de manutenção.

1.3 Escopo

O escopo deste trabalho se resume no desenvolvimento de protótipos de sistemas de diagnóstico baseado em técnicas de inteligência computacional, com a finalidade de auxiliar na detecção de faltas em transformadores de potência, bem como auxiliar o setor de manutenção, na programação e planejamento das paradas da planta, no sentido de reduzir as perdas na produção em consequência das paradas não programadas, associadas às faltas nestes equipamentos.

1.4 Objetivos

O objetivo geral deste trabalho é desenvolver um sistema baseado na aplicação de técnicas de Inteligência computacional que terá por finalidade diagnosticar faltas incipientes em trans-

formadores de potência através da análise dos gases que se encontram dissolvidos no óleo isolante, tal como o apresentado em (HELL et. al, 2002a).

1.5 Metodologia

O protótipo de sistema de diagnóstico foi construído através da aplicação das técnicas de Inteligência Computacional, tais como Rede Neural de Múltiplas Camadas (MLP²), Rede Neural de Função de Base Radial (RBF³), Sistema de Inferência Fuzzy Neuro-Adaptativa (ANFIS⁴) e Rede Neural de Kohonen (SOM⁵). O treinamento dos sistemas de diagnóstico foi feito com base em dados de análises cromatográficas fornecidos pela Companhia Energética de Minas Gerais (CEMIG), sendo que os diagnósticos associados à estas análises foram obtidos aplicando-se a norma internacional (ANSI/IEEE Std C57.104- 1991) referente à interpretação de gases dissolvidos em óleo isolante. A partir dos resultados de diagnóstico obtidos, construiu-se a base de dados utilizada no processo de treinamento do sistema desenvolvido.

Com isso, o trabalho foi dividido em cinco fases, ordenadas cronologicamente a seguir:

Primeira Fase : Esta fase consistiu no estudo das normas internacionais e nacionais existentes relativas aos procedimentos de ensaios e indicativos de faltas em transformadores, bem como de literatura técnica específica sobre métodos de diagnóstico de faltas, os quais farão parte do estado da arte do trabalho;

Segunda Fase : Esta fase foi subdividida em duas partes, uma onde foi realizado um estudo sobre técnicas inteligência computacional e outra na qual foram ser obtidos e analisados dados relativos aos ensaios realizados em transformadores de potência fornecidos pela Companhia Energética de Minas Gerais (CEMIG), com a finalidade da obtenção da assinatura de defeitos com relação as condições dos transformadores;

Terceira Fase : Nesta fase foram desenvolvidos diversos protótipos de sistemas de diagnóstico baseados na análise de gases dissolvidos usando como ferramentas redes neurais , lógica fuzzy e sistemas híbridos associando redes neurais com lógica fuzzy, com a finalidade de

²Do inglês *Multi Layer Perceptron*

³Do inglês *Radial Basis Function*

⁴Do inglês *Adaptative Neuro Fuzzy Inference System*

⁵ Do inglês *Self Organizing Map*

se obter as condições de faltas e diagnóstico dos transformadores, com base nos dados obtidos na Segunda Fase;

Quarta Fase : Consiste no teste e validação dos protótipos desenvolvidos na fase 3;

Quinta Fase : Redação do relatório final da pesquisa na forma de dissertação;

1.6 Estrutura do Trabalho

Este trabalho esta organizado da seguinte forma:

- No Capítulo 1 foi feita uma breve introdução do problema proposto, bem como a apresentação da relevância, o escopo, os objetivos e a metodologia deste trabalho;
- No Capítulo 2 é apresentado a revisão das técnicas de diagnóstico de faltas em transformadores de potência e das técnicas de inteligência computacional aplicadas a este problema;
- No Capítulo 3 é apresentada todos os fundamentos teóricos que serviram de base para a construção do protótipo de sistema de diagnóstico desenvolvido;
- O Capítulo 4 apresenta os sistemas desenvolvidos com seus respectivos resultados experimentais;
- No Capítulo 5 são apresentadas as conclusões do trabalho, bem como a proposição de trabalhos futuros.

Capítulo 2

Revisão de Literatura

2.1 Diagnóstico de Falhas em Transformadores

*D*urante a operação de um transformador, o óleo isolante e os outros materiais dielétricos sofrem, sob a ação de temperatura e de tensões elétricas, processos de decomposição química que resultam na evolução de gases. A ocorrência de falhas, tais como superaquecimento, arco ou descargas parciais, resultam na produção de quantidades apreciáveis de gases, que são tradicionalmente utilizados para acionar o relê Buchholz, que atua como alarme e proteção dos transformadores.

A possibilidade de esclarecer a natureza ou a gravidade da falta em caso de operação do relê Buchholz, vem sendo considerada desde 1956 (HOWE et. al., 1956) por meio da análise do gás coletado.

Estudos realizados ao longo do tempo indicaram ser possível relacionar a existência de determinados gases com a natureza da falta que os geraram, ou com o material por esta atingido. Deste modo, o estudo dos gases recolhidos no relê Buchholz (HOWE et. al., 1956) possibilita distinguir um possível alarme falso, e ainda fornece informações quanto ao tipo de material isolante envolvido na falta, tipo da falta, sua severidade e localização.

Os gases formados pela decomposição dos materiais isolantes são total ou parcialmente dissolvidos no óleo, sendo diluídos e transportados de todos os pontos. Isto possibilita, através da coleta de uma amostra, a obtenção de informações de todas as partes em contato com o óleo.

Os métodos antiquados para detectar gases combustíveis, baseados em ensaios de flama- bilidade ou análise química, não representam grande sensibilidade, e são satisfatórios apenas

em casos de deterioração avançada do óleo ou papel isolantes.

O advento de técnicas de análise de gás, dentre as quais se destaca a cromatografia, capazes de processar pequenas amostras com grande sensibilidade e precisão, tornou possível um novo enfoque do problema. No caso da ocorrência de uma falta incipiente, as quantidades de gases gerados são pequenas. Estes gases dissolvem-se no óleo, podendo não resultar em uma fase gasosa capaz de ser detectada ou examinada como anteriormente descrito. A extração e análise de uma amostra do óleo isolante constitui, no entanto, um instrumento poderoso para a identificação precoce de uma falta.

Esse método foi desenvolvido e é largamente utilizado em todo o mundo, permitindo um diagnóstico das condições internas de um transformador, segundo vários critérios descritos posteriormente, na Seção 2.2.

2.1.1 Tipos de faltas e suas causas

Segundo (SPUDAT, 1993), é possível dividir as faltas operativas que provocam a geração de gases em basicamente dois grupos: o das faltas de origem elétrica e o das faltas de origem térmica:

Faltas Elétricas

Descargas parciais de baixa intensidade: Aparecem devido à baixa eficiência dos materiais isolantes ou existência de gás na isolação. Esta falta não gera sinais visíveis de deterioração dos materiais envolvidos;

Descargas parciais de alta intensidade: Aparecem devido à baixa eficiência dos materiais isolantes, existência de gás na isolação ou sobre-solicitação do isolamento. Diferentemente das descargas parciais de baixa intensidade, esta falta gera sinais visíveis de deterioração dos materiais envolvidos;

Descargas de baixa energia: Ocorrem com o mau contato de conexões de potenciais, faiscamento em rupturas de condutores ou descargas descontinuadas em contatos de seletores de comutação de tensão, devido à rugosidades geradas nas superfícies dos contatos, geradas pelo continuado processo de comutação sobre carga;

Descargas de alta energia: Ocorrem devido a curto-circuitos internos dos enrolamentos, curto-circuitos em conexões de potenciais contra terra ou descargas entre condutores não isolados;

Faltas Térmicas

Sobreaquecimento de 150 ° C até no máximo 300 ° C: Surgem devido ao sobreaquecimento dos enrolamentos ou pontos quentes no núcleo, ocasionado por perdas excessivas de magnetização, refrigeração insuficiente ou sobrecargas;

Sobreaquecimento local de 300 ° C até no máximo 1000 ° C: Ocorrem devido a correntes circulantes (correntes de Foucault) no núcleo, tanque, armação ou ainda devido ao mau contato em seletores de tensão, devido às rugosidades geradas nas superfícies dos contatos, geradas pelo continuado processo de comutação sobre carga, com carbonização do óleo isolante;

Sobreaquecimento local acima de 1000 ° C: Ocasionado por curto-circuitos interlaminares no núcleo com pontos de fusão ou condutores fundidos e pontos metálicos com início de fusão.

Cada uma das faltas acima mencionadas podem ocorrer isoladamente ou associadas entre si.

2.1.2 Gases Normalmente Encontrados Dissolvidos no Óleo Isolante

Segundo a norma (ANSI/IEEE Std C57.104- 1991), os gases que são normalmente encontrados dissolvidos no óleo isolante são os listados na Tabela 2.1.

Com base em seus valores absolutos, relação existente entre gases e taxa de crescimento é possível avaliar as condições operativas do equipamento. Para tanto é necessário conhecer um pouco do mecanismo de decomposição do óleo isolante e dos isolantes celulósicos, normalmente utilizados nos transformadores.

Tabela 2.1: Gases Normalmente encontrados no óleo

Hidrogênio	H_2
Oxigênio	O_2
Nitrogênio	N_2
Metano	CH_4
Monóxido de Carbono	CO
Dióxido de Carbono	CO_2
Etileno	C_2H_4
Etano	C_2H_6
Acetileno	C_2H_2
Propano	C_3H_8
Propileno	C_3H_6

2.1.3 Decomposição do Óleo Isolante e dos Materiais Celulósicos

Ocorrendo uma anormalidade, ou falta, em um transformador, ocorre também uma decomposição do óleo isolante ou da celulose. Esta decomposição depende do fator gerador da anormalidade, conforme mostrado em (SPUDAT, 1993) e resumido a seguir:

Decomposição do Óleo Isolante

As descargas elétricas causam principalmente a pirólise¹ no óleo. Com isso, além do carbono que se dispersa no óleo em forma de fuligem, surgem os gases oriundos da dissociação, ou seja, hidrogênio e acetileno. A energia necessária para dissociar estes gases vai depender da origem do óleo isolante. Um óleo com maior teor de aromáticos dissocia menor quantidade de gás.

Na pirólise, os hidrocarbonetos aromáticos são dissociados em quantidades apreciáveis em temperaturas superiores a 900 ° C. As alterações de estrutura molecular que ocorrem durante a pirólise são dos seguinte tipos:

- Reação de Desidratação : Quebra de ligações $C - -H$;
- Reação de Decomposição : Quebra de ligações $C - -C$;
- Reação de Constituição : Polimerização e condensação .

A decomposição do óleo a temperaturas inferiores a 400 ° C é pequena. Porém quando

¹processo de decomposição química por calor na ausência de oxigênio, em que a matéria orgânica pode ser convertida em diversos subprodutos

submetido a solicitações prolongadas, surgem apreciáveis quantidades de gases combustíveis, principalmente hidrocarbonetos saturados (metano, etano, propano).

Acima de 400 ° C a decomposição do óleo é muito rápida e surgem, principalmente, hidrocarbonetos insaturados (etileno e propileno).

Decomposição dos Materiais Celulósicos

A decomposição térmica de celulose produz essencialmente monóxido de carbono (CO) e dióxido de carbono (CO_2).

2.2 Critérios de Análise de Gases Dissolvidos

Tentativas de diagnosticar faltas em transformadores a partir de gases gerados após a ocorrência de faltas tiveram início na década de 50, tendo como base os gases coletados no relé Buchholz. Em 1956 HOWE publicou (HOWE et. al., 1956) uma avaliação detalhada de faltas a partir dos gases coletados no relé. Embora a importância da análise dos gases coletados no relé Buchholz fosse inquestionável, os resultados obtidos eram geralmente tardios, pois para se ter uma quantidade considerável de gases que permitissem realizar o diagnóstico, a degradação interna no transformador já havia atingido estágio avançado. Apenas com a chegada de técnicas de cromatografia líquida, capazes de analisar pequenas amostras de óleo com grande precisão e sensibilidade, é que foi possível ter uma nova visão do problema. Em 1968 foi iniciado um acompanhamento regular de gases dissolvidos no óleo utilizando-se análise cromatográfica e, segundo Halstead (HALSTEAD, 1973), em 1970 foram monitorados mais de 1.000 transformadores com faixas de tensão que variavam entre 132, 275 e 400 KV. Os dados coletados mostraram que todos os transformadores, incluindo os levemente carregados, desenvolviam hidrogênio e outros hidrocarbonetos gasosos, embora em pouca quantidade. Em 1973 Halstead publicou um estudo teórico com base na avaliação termodinâmica dos óleos isolantes (HALSTEAD, 1973) no qual sugeriu que a produção de cada hidrocarboneto no óleo varia de acordo com a temperatura do seu ponto de decomposição. Isto conduziu à hipótese de que cada gás alcançaria o seu máximo grau de concentração em uma temperatura específica. A partir desse estudo, diversos métodos de diagnóstico de faltas a partir da análise de gases dissolvidos no óleo isolante foram propostos, dentre os quais pode-se destacar:

2.2.1 Método dos Gases Chave

O diagnóstico através do método dos Gases Chave baseia-se na predominância de determinado gás com relação ao Total de Gases Combustíveis Dissolvidos (TDCG²) no óleo isolante. O TDCG é calculado somando-se as concentrações de hidrogênio (H_2), metano (CH_4), etano (C_2H_6), etileno (C_2H_4), acetileno (C_2H_2) e monóxido de carbono (CO) que encontram-se dissolvidos no óleo. Nesse método as concentrações absolutas (em ppm) e as taxas de geração (em ppm/dia) dos gases são usadas para se determinar o tipo e a intensidade de determinadas faltas. Essas faltas estão relacionadas com os "gases chave" da seguinte forma:

H_2	⇒	Corona (falta elétrica)
CH_4 & C_2H_6	⇒	Degradação do óleo a baixas Temperaturas (falta térmica)
C_2H_4	⇒	Degradação do óleo a altas Temperaturas (falta térmica)
C_2H_2	⇒	Arco (falta elétrica)
CO & CO_2	⇒	Degradação de isolantes celulósicos (relacionado ao processo de envelhecimento)

Nos transformadores em operação normal o óleo isolante sempre contém consideráveis concentrações de dióxido e monóxido de carbono, certa quantidade de hidrogênio e pequenas concentrações de hidrocarbonetos leves. Esses gases são gerados, em até certos níveis limite, devido ao processo natural de envelhecimento dos materiais isolantes. Se, para cada gás, pudessem ser atribuídos limites para esses níveis de concentrações, os quais uma vez excedidos em pequenas proporções (5 ou 10%) servissem como primeiro indicativo de uma anormalidade, conduzindo, assim, a exames mais detalhados, tais níveis poderiam ser considerados como "normais". Esses valores não devem ser generalizados para os diversos tipos de transformadores existentes, pois transformadores mais velhos, por exemplo, mesmo aparentemente livres de faltas possuem, altas concentrações de gases, já nos mais novos essas concentrações são baixas. A magnitude dessas concentrações normais dependem largamente de fatores como idade, condições de operação, mas valores limites podem ser estabelecidos empiricamente.

Dörnenburg et. al. (DÖRNENBURG e STRITTMATTER, 1974) baseado em dados experimentais identificou que se as concentrações de gases apresentadas na Tabela 2.2 não forem excedidas após poucos anos de operação dos transformadores, então elas indicam que eles estão isentos de falta.

² Do inglês *Total Dissolved Combustible Gas*

Tabela 2.2: Limites de Concentração

Hidrogênio	200 ppm
Metano	50 ppm
Etano	15 ppm
Etileno	60 ppm
Acetileno	15 ppm
Monóxido de Carbono	1000 ppm
Dióxido de Carbono	11000 ppm

Se dois ou mais gases excederem as concentrações limite mostradas, o transformador deve ser considerado como suspeito.

2.2.2 Método das Razões de Dörnenburg

Em 1970, Fallou et. al. (FALLOU et al., 1970) diferenciaram faltas de origem elétricas de faltas de origem térmicas pela comparação de pares de gases característicos, os quais possuíam coeficientes de solubilidade e de difusão próximos. Esse método foi considerado bastante promissor, pois eliminava o efeito do volume de óleo do transformador, podendo ser aplicado tanto à unidades com grandes volumes de gases gerados como à pequenas unidades. A partir dessas considerações foram obtidas 4 relações de concentração de pares de gases particularmente úteis (DÖRNENBURG e STRITTMATTER, 1974), que são apresentados a seguir:

1. $\frac{CH_4}{H_2} \Rightarrow \frac{\text{Concentração de Metano}}{\text{Concentração de Hidrogênio}}$
2. $\frac{C_2H_2}{C_2H_4} \Rightarrow \frac{\text{Concentração de Acetileno}}{\text{Concentração de Etileno}}$
3. $\frac{C_2H_6}{C_2H_2} \Rightarrow \frac{\text{Concentração de Etano}}{\text{Concentração de Acetileno}}$
4. $\frac{C_2H_2}{CH_4} \Rightarrow \frac{\text{Concentração de Acetileno}}{\text{Concentração de Metano}}$

A partir de dados experimentais obteve-se uma relação entre faltas típicas e faixas de valores das relações anteriores, gerando assim um critério para o diagnóstico de faltas. Essas faixas, tanto para gases livres (tais como os contidos no relê Buchholz) quanto para gases dissolvidos no óleo, são mostrados nas Tabelas 2.3 e 2.4, respectivamente.

O fluxograma apresentado na figura 2.1 ilustra passo a passo a aplicação do método das

Tabela 2.3: Faixas de Valores Característicos das Relações de Gases Dissolvidos no óleo

Relações de Concentrações de Gases Dissolvidos	$\frac{CH_4}{H_2}$	$\frac{C_2H_2}{C_2H_4}$	$\frac{C_2H_6}{C_2H_2}$	$\frac{C_2H_2}{CH_4}$
Tipos de Falhas Características				
Sobreaquecimento local	> 1.0	< 0.75	> 0.4	< 0.3
Descargas de baixa Intensidade	< 0.1	*	> 0.4	< 0.3
Todos os outros tipos de descargas	>0.1 e <1	> 0.75	< 0.4	> 0.3

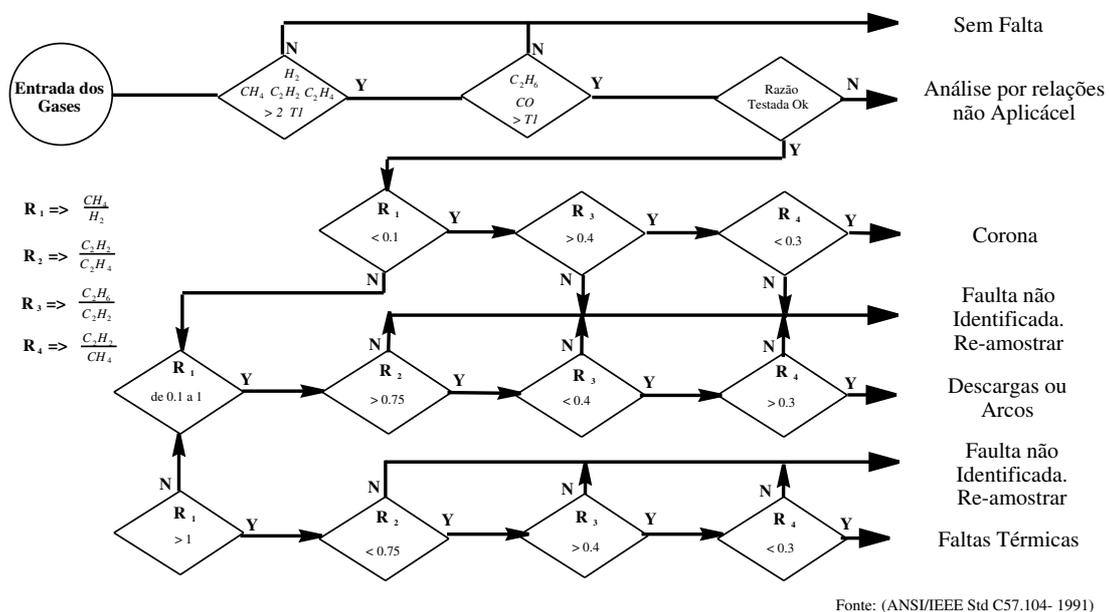
* Não indica o tipo de falta, geralmente não aplicável Fonte: (ANSI/IEEE Std C57.104- 1991)

Tabela 2.4: Faixas de Valores Característicos das Relações de Gases Livres (relé ou almofada)

Relações de Concentrações de Gases Dissolvidos	$\frac{CH_4}{H_2}$	$\frac{C_2H_2}{C_2H_4}$	$\frac{C_2H_6}{C_2H_2}$	$\frac{C_2H_2}{CH_4}$
Tipos de Falhas Características				
Sobreaquecimento local	> 0.1	<1.0	> 0.2	< 0.1
Descargas de baixa Intensidade	< 0.01	*	> 0.2	< 0.1
Todos os outros tipos de descargas	> 0.01 e < 0.1	>1.0	< 0.2	> 0.1

* Não indica o tipo de falta, geralmente não aplicável Fonte: (ANSI/IEEE Std C57.104- 1991)

razões de Dörnenburg.



Fonte: (ANSI/IEEE Std C57.104- 1991)

Figura 2.1: Fluxograma de Aplicação do Método das Razões de Dörnenburg

A seguir são descritos os passos indicados na Figura 2.1:

Passo 1 : As concentrações dos gases são obtidas através da análise cromatográfica do óleo

isolante;

Passo 2 : Se ao menos uma das concentrações avaliadas para H_2 , CH_4 , C_2H_2 e C_2H_4 exceder duas vezes os valores limites apresentados na Tabela 2.2, e uma das outras três concentrações exceder o valor limite da Tabela 2.2, o equipamento testado é considerado defeituoso; deve-se, então, prosseguir para o Passo 3 para verificar a validade do método;

Passo 3 : Determinando a validade do procedimento : Se ao menos um dos gases em *cada uma das relações* apresentadas nas Tabelas 2.3 e 2.4 exceder o valor limite da Tabela 2.2, o método é válido; caso contrário, as relações não são significantes e o equipamento deve ser reamostrado e investigado por métodos alternativos, tais como o método de avaliação do transformador usando concentrações individuais e o TDCG e o método de avaliação do possível tipo de falta pelo método dos gases chave, descritos em (ANSI/IEEE Std C57.104- 1991);

Passo 4 : Assumindo que o método de análise é válido, cada relação é comparada com os valores limites obtidos a partir das Tabelas 2.3 e 2.4;

Passo 5 : Se todas as relações estão dentro dos limites apresentados nas Tabelas 2.3 e 2.4 para um tipo de falta específico, então o tipo de falta sugerido é válido.

Esse método de diagnóstico deve ser usado apenas com extrema precaução pois se os gases dissolvidos forem originados a partir de faltas que podem não persistir ao longo do tempo, estes gases de decomposição se desprendem para a superfície do óleo, expandindo-se no tanque do transformador e perdendo-se na atmosfera. Este fato pode causar distorções no diagnóstico.

Monóxido e dióxido de carbono são tipicamente relacionados ao processo de decomposição do isolamento sólido, e não são usados nas relações características.

2.2.3 Método das Razões de Rogers

Em 1975 Rogers (ROGERS, 1975) produziu um refinado código a partir das faixas de relações de gases, utilizando-se as quatro relações mostradas anteriormente, porém diagnosticando um maior número de faltas . O uso do código facilita a programação computacional no desenvolvimento de sistemas para diagnóstico de falta.

Posteriormente, com base em estudos estatísticos em mais de dez mil análises de gás em óleo isolante de transformadores, Rogers (ROGERS, 1978) mostrou que certos tipos de condições de faltas poderiam ser diferenciados em faixas mais detalhadas a partir de combinações de relações de gases. Isso foi confirmado pelo exame interno de um certo número de transformadores suspeitos juntamente com unidades destruídas em faltas, bem como pelo estudo de "pontos quentes" prováveis de ocorrer em transformadores sob condições operacionais.

Com a finalidade de se estabelecer a identificação de faltas reais Rogers (ROGERS, 1978) realizou um estudo em cem conjuntos de análises de óleo extraídos de transformadores com conhecidos tipos de falta a fim de avaliar a provável temperatura na qual as relações indicam mudanças significativas. Com base no resultado desse estudo e de avaliações teóricas, novos valores das relações para faltas elétricas e térmicas foram então obtidas. Assim, pôde-se estabelecer que o uso de apenas três relações simplificaria a interpretação. Para ajudar no entendimento da técnica as tabelas anteriores foram reorganizadas para indicar uma progressão mais racional de faltas, resultando no código descrito na Tabela 2.5.

Tabela 2.5: Código para Análise de Gases Dissolvidos em Óleo Mineral

Código de Faixas de Relação	Razão de Gases Característicos		
	$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$
< 0.1	0 (L)	1 (L)	0 (L)
0.1 - 1.0	1 (M)	0 (M)	0 (L)
1.0 - 3.0	1 (M)	2 (H)	1 (M)
> 3.0	2 (H)	2 (H)	2 (H)
Código de Falta	Tipo de Falta		Característica
0	Operação Normal		
	Falta Térmica de Baixa		
1	Intensidade < 150°		0 (L) 0 (M) 1 (M)
	Falta Térmica de Baixa		
2	Intensidade 150° - 300°		0 (L) 2 (H) 0 (L)
	alta Térmica de Média		
3	Intensidade 300° - 700°		0 (L) 2 (H) 1 (M)
	Falta Térmica de Alta		
4	Intensidade > 700°		0 (L) 2 (H) 2 (H)
	Descargas Parciais de		
5	Baixa Intensidade		0 (L) 1 (L) 0 (L)
	Descargas Parciais de		
6	Alta Intensidade		1 (M) 1 (L) 0 (L)
7	Descargas de Baixa Intensidade		1-2 (M-H) 0 (M) 1-2 (M-H)
8	Descargas de Alta Intensidade		1 (M) 0 (M) 2 (H)

Esse código foi incluído no documento IEC 10A 53 e passou a ser recomendado pelo IEEE e IEC como principal código para interpretação de faltas incipientes em transformadores, usando análise de gases no óleo isolante.

2.3 Ferramentas Para Diagnóstico de Faltas a Partir dos Critérios DGA

Nesta seção são apresentadas recentes ferramentas baseadas em princípios de inteligência computacional, tais como sistemas de inferência fuzzy e redes neurais, que foram propostas com a finalidade de diagnosticar faltas incipientes em transformadores e aumentar a precisão do método de gases dissolvidos (DGA). Dentre essas técnicas, as mais usadas são:

2.3.1 Lógica Fuzzy

Em 1993 Dukarm (DUKARM, 1993) verificou que uma abordagem utilizando técnicas fuzzy e sistemas de inferência fuzzy aplicariam-se aos método DGA apresentados anteriormente . Comum à esses três métodos é o uso de limiares numéricos para classificar as características dos gases dissolvidos em vários intervalos. As informações contidas nos intervalos poderiam ser usadas para deduzir um diagnóstico. Segundo Dukarm, esta abordagem seria utilizada da seguinte forma:

Método dos Gases Chave com lógica fuzzy

O método dos gases chave é bem apropriado para servir de base para a implementação de um sistema de inferência fuzzy. Para tanto, um padrão de gases chave, bem como os limites destes gases, devem ser definidos para cada transformador ou para grupos de transformadores similares.

A partir destes limites são definidos conjuntos fuzzy para o Total de Gases Combustíveis Dissolvidos (TDCG), para cada gás combustível, para a relação de cada gás combustível com o TDCG e para a taxa de evolução dos gases em relação à amostra anterior, da seguinte forma:

- Para o TDCG e para cada gás combustível são associados os conjuntos "Normal", "Alto" e "Muito Alto" referentes aos seus respectivos níveis de concentração no óleo;

- Para cada gás combustível são associados os conjuntos "Alto" e "Muito Alto" referentes aos níveis dos gases em relação ao TDCG;
- Para o TDCG e para cada gás combustível são associados os conjuntos "Alto" e "Muito Alto" referentes às taxas de geração dos gases.

A partir desses conjuntos são definidas regras fuzzy, onde cada regra mostra como um diagnóstico é derivado a partir da classificação de dados de entrada.

Assim, este sistema pode ser resumido da seguinte forma:

Entrada de Dados: Valores Reais associados a cada uma das variáveis de entrada (valor do TDCG, concentração cada gás combustível, relação da concentração de cada gás combustível com o TDCG e a taxa de evolução dos gases) são fornecidos ao sistema;

Fuzificação: Os dados fornecidos são aplicados nos conjuntos descritos anteriormente, gerando um número fuzzy (ver Seção 3.1);

Inferência: A partir das entradas fuzificadas, todas as regras são verificadas e para cada regra é associado um grau de ativação. Por Exemplo:

$$\begin{array}{l} \text{SE } \text{TDCG é "Normal"} \text{ E } \%C_2H_2 \text{ é "Alto"} \\ \text{ENTÃO } \text{Falta é "Sobreaquecimento"} \end{array}$$

onde $\%C_2H_2$ é o percentual de Acetileno em relação ao TDCG.

O grau de ativação da regra é obtido tomando-se o mínimo índice de pertinência das variáveis de entrada nos seus respectivos conjuntos (Modelo fuzzy de Mamdani - Seção 3.2.1). Desta forma, se no exemplo anterior os graus de ativação apresentados pelo TDCG no conjunto "Normal" e pelo $\%C_2H_2$ no conjunto "Alto" fossem respectivamente 0.8 e 0.6, o grau de ativação da regra, ou seja, o índice de certeza de que a falta associado às variáveis de entrada é "Sobreaquecimento", seria igual a 0.6;

Diagnóstico: Após todas as regras terem sido avaliadas, o diagnóstico final é inferido como sendo aquele associado à regra com maior grau de ativação.

Método das Razões de Rogers com lógica fuzzy

O método das relações de Rogers, como descrito anteriormente, usa três razões de gases, obtidas através de análises cromatográficas.

No modelo fuzzy um conjunto fuzzy é associado a cada uma das relações de entrada, com valores de limiar definidos de acordo com o método original, conforme pode ser visto nas figuras apresentadas a seguir, onde a Figura 2.2 mostra a relação $\frac{CH_4}{H_2}$, a Figura 2.3 mostra a relação $\frac{C_2H_2}{C_2H_4}$ e a Figura 2.4 mostra a relação $\frac{C_2H_4}{C_2H_6}$.

As regras fuzzy são derivadas do código apresentado na Tabela 2.5 e são aplicadas para obter os graus de ativação para o diagnóstico baseado no método de Rogers.

Com a finalidade de assegurar a consistência com o método de Rogers, as funções de pertinência fuzzy foram definidas de tal forma que nos limites dos correspondentes intervalos reais o índice de pertinência de ambos os conjuntos fuzzy ativos nestes pontos é igual a 0.5 (DUKARM, 1993). Isso assegura que o grau de ativação do diagnóstico será maior do que 0.5 se e somente se o método padrão de Rogers produz um diagnóstico.

Para a obtenção de um diagnóstico, o valor das relações de gases são aplicados em um sistema de inferência fuzzy, onde as funções de pertinência associadas aos conjuntos fuzzy (Figuras 2.2 , 2.3 , 2.4) são associadas às variáveis consideradas.

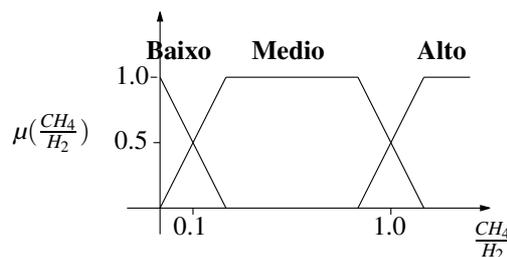


Figura 2.2: Relação Metano / Hidrogênio

Da mesma forma que para o método dos gases chave, as regras são testadas, conforme mostra a Figura 2.5, para uma dada entrada, sendo que a regra que possui maior grau de ativação será a regra associada ao diagnóstico final. A saída do sistema fuzzy será um dos tipos de falta apresentados na Tabela 2.5, onde L, M e H, são os conjuntos fuzzy **Baixo**,

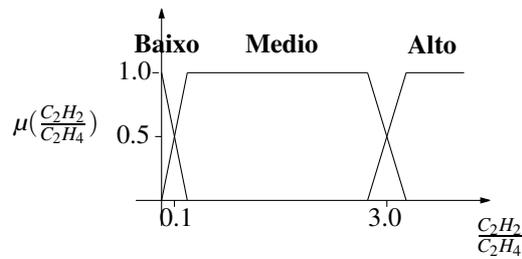


Figura 2.3: Relação Acetileno / Etileno

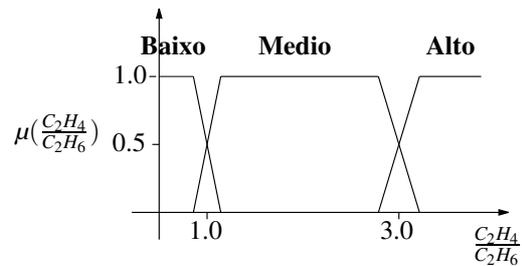


Figura 2.4: Relação Etileno / Etano

Médio e Alto, respectivamente.

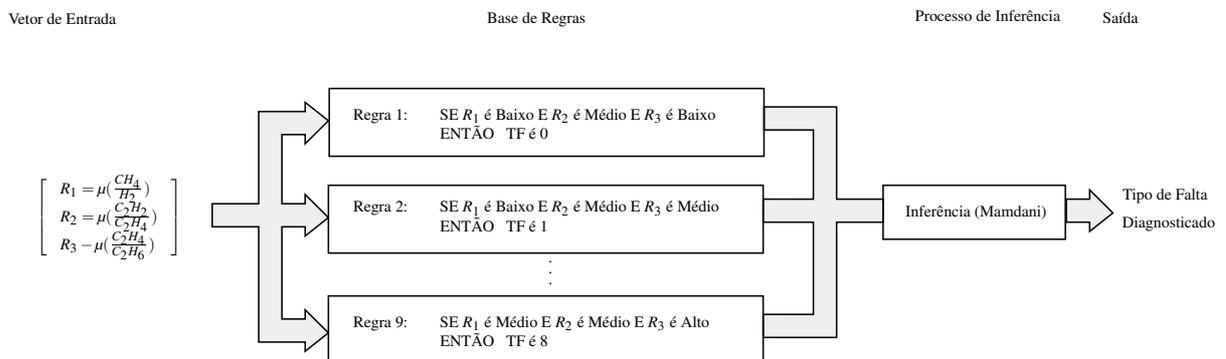


Figura 2.5: Sistema de Inferência Fuzzy para diagnóstico de falhas em transformadores baseado no método de Rogers

Um tratamento similar pode ser aplicado ao método das relações de Dörnenburg.

2.3.2 Redes Neurais Artificiais

Redes Neurais Artificiais (RNA's) também tem sido utilizadas com a finalidade de diagnosticar falhas incipientes em transformadores através da análise de gases dissolvidos no óleo isolante. Sua aplicação é baseada no fato de que uma RNA, por intermédio de processos de treinamento,

pode revelar mecanismos complexos que podem ser desconhecidos.

Para tanto é necessário uma criteriosa seleção dos parâmetros da RNA a ser utilizada, tais como topologia da rede, número de camadas intermediárias, número de neurônios por camada, bem como as características de entrada e saída. Esses parâmetros podem ser obtidos experimentalmente, por meio de processos tentativa e erro, a fim de determinar o melhor número de camadas intermediárias.

Zang et. al. (1996) mostraram que uma abordagem "dois-passos" poderia alcançar melhores performances no diagnóstico, sendo útil, também, para distinguir faltas relativas à decomposição da celulose de faltas relativas á decomposição do óleo.

Esta abordagem consiste na construção de duas RNA's distintas. A primeira tem a finalidade de diagnosticar os principais tipos de falta no transformador, tais como sobreaquecimento, corona e arcos, incluindo a condição normal. A segunda RNA tem a finalidade de diagnosticar apenas a degradação da celulose. Desta forma é possível diagnosticar sete possíveis padrões de falta (incluindo a condição normal), uma vez que cada um dos principais tipos de falta pode ou não estar acompanhado de processos de degradação da celulose. As duas RNA's propostas por Zhang podem ser vistas nas Figuras 2.6 e 2.7, respectivamente.

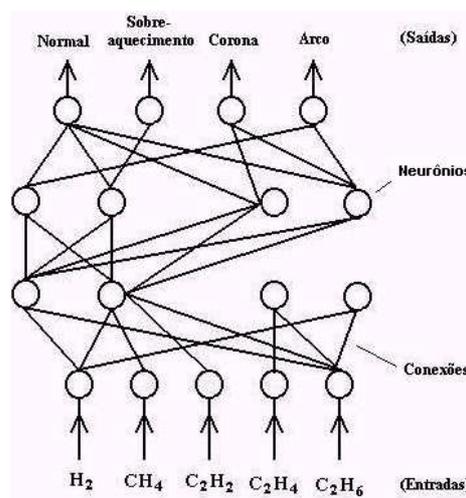


Figura 2.6: RNA para diagnóstico dos principais tipos de falta proposta por Zhang

Estudos mais recentes (WANG et. al., 2000) propõem a construção de uma RNA para o diagnóstico de cada um dos principais tipos de falta, conforme Figura 2.8, onde temos

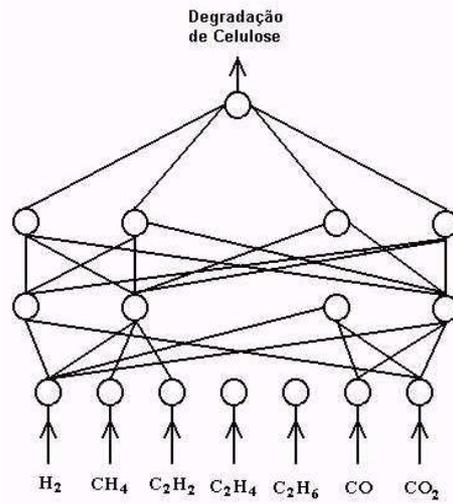


Figura 2.7: RNA para diagnóstico de degradação da celulose proposta por Zhang

a seguinte notação: OH (sobreaquecimento do óleo ou celulose), OHO (sobreaquecimento do óleo), LED (descargas parciais de baixa intensidade), HEAD (descargas parciais de alta intensidade ou arco) e CD (degradação da celulose).

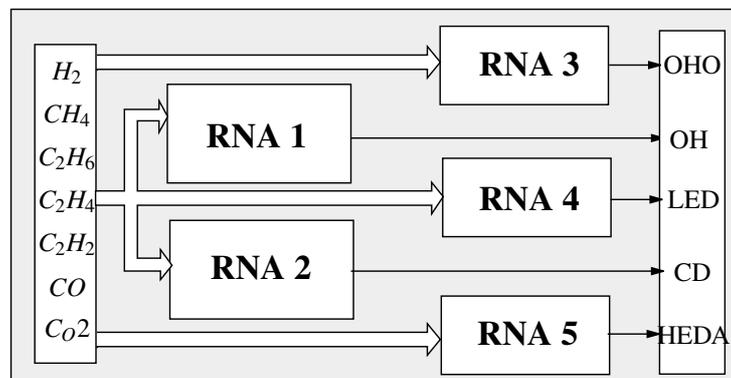


Figura 2.8: Sistemas de RNA's proposta por Wang

Em ambas propostas foram obtidas grande precisão no diagnóstico, porém, para efeitos reais, essa precisão depende, em grande parte, da consistência e do volume dos dados de treinamento.

Capítulo 3

Fundamentação Teórica

A base teórica necessária para o desenvolvimento e implementação dos protótipos de sistema de diagnóstico desenvolvidos no Capítulo 4 desta dissertação é aqui apresentada. Com a finalidade de se tratar de forma adequada os dados a serem utilizados nos sistemas de diagnóstico propostos, desenvolveu-se inicialmente sistemas de classificação. Estes sistemas foram aplicados a uma base de dados reais contendo resultados de análises cromatográficas em amostras de óleo isolante coletados em transformadores de potência. Posteriormente, foram analisadas algumas ferramentas de inteligência computacional que serviriam de base para o desenvolvimento dos sistemas de diagnóstico implementados, tais como: rede neural de múltiplas camadas treinada com algoritmo *backpropagation*, rede neural de Kohonen, rede neural de Função Base Radial (RBFN), sistema de inferência fuzzy (FIS) e sistema de inferência fuzzy neuro-adaptativo (ANFIS).

3.1 Teoria de Conjuntos Fuzzy

O conceito de conjunto Fuzzy foi introduzido, em 1965, por Lotfi A. Zadeh (Universidade da Califórnia, Berkeley) com a finalidade de solucionar problemas de natureza industrial, biológica ou química, que compreendessem situações imprecisas, não passíveis de processamento através da lógica computacional fundamentada na lógica booleana.

Baseado nesse conceito foi desenvolvida a Lógica Fuzzy, que fornece os princípios formais de raciocínio aproximado, com raciocínio preciso visto como um caso específico (ZADEH, 1988).

Em sistemas baseados em lógica fuzzy, ao contrário de sistemas lógicos clássicos, procura-se modelar a forma imprecisa de raciocínio, que desempenha um papel essencial na notável capacidade humana de tomar decisões em ambientes de incerteza e imprecisão.

O desenvolvimento e aplicação de lógica fuzzy foram motivados pelas duas principais razões pelas quais sistemas lógicos convencionais não podem ser empregados na solução de questões baseadas em conhecimento inexato, incompleto ou não totalmente preciso: 1) Estes sistemas não fornecem uma representação do conhecimento de proposições expressas em uma linguagem natural quando este conhecimento é impreciso; e 2) Nos casos em que o conhecimento pode ser representado de forma simbólica em uma linguagem de representação de conhecimento, como por exemplo redes semânticas ou gráficos de dependência conceitual, não existem mecanismos que possam inferir uma resposta.

Com a aplicação de lógica fuzzy é possível tratar esses problemas da seguinte forma: 1) O conhecimento de uma proposição imprecisa é representada como uma restrição flexível de uma ou mais variáveis; e 2) a resposta para a questão proposta é deduzida através da propagação de restrições flexíveis.

Nas seções seguintes serão apresentados os conceitos fundamentais de Conjuntos Fuzzy e de Lógica Fuzzy, assim como algumas definições e operações que permitem tratar os mecanismos de inferência, base para o que se convencionou chamar de Raciocínio Aproximado.

3.1.1 Conjuntos Fuzzy

Na teoria clássica dos conjuntos, um elemento de um universo pertence ou não a um determinado conjunto. Assim, pode-se definir a pertinência de um elemento particular x de um conjunto A em um determinado universo \mathcal{U} através da função característica:

$$\mu_A(x) : \mathcal{U} \rightarrow \{0, 1\} \quad \mu_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{caso contrário} \end{cases} \quad (3.1)$$

Na teoria de conjuntos fuzzy é proposta uma caracterização mais ampla, generalizando a noção de função característica de modo que ela possa assumir valores no intervalo $[0,1]$, ou seja, um **conjunto fuzzy** A em um **universo** \mathcal{U} é definido como sendo um conjunto de pares ordenados da forma:

$$A = \{(x, \mu_A(x)) | x \in \mathcal{U}\}, \quad (3.2)$$

onde $\mu_A(x)$ é chamada de **função de pertinência** (MF¹) do conjunto A . A MF associa a cada um dos elementos de \mathcal{U} um grau de pertinência (ou valor de pertinência) entre 0 e 1.

O conjunto suporte de um conjunto fuzzy A é o subconjunto de pontos x de \mathcal{U} tal que $\mu_A(x) > 0$. Um conjunto fuzzy cujo conjunto suporte é um único ponto de \mathcal{U} com $\mu_A = 1$ é chamado de conjunto unitário fuzzy ou *singleton*.

3.1.2 Definições e Operações

União, interseção e complemento são as mais básicas operações na teoria de conjuntos clássicos. Com base nestas operações foram definidas operações similares para conjuntos fuzzy. Porém, antes da introdução destas três operações, é necessário definir inicialmente a noção de subconjunto fuzzy, que desempenha um importante papel, tanto na teoria de conjuntos clássicos quanto para conjuntos fuzzy. A definição de subconjunto fuzzy é uma extensão natural de subconjuntos clássicos, e pode ser definida como:

Um conjunto fuzzy A é um subconjunto de um conjunto fuzzy B se e apenas se :

¹ do inglês *Membership Function*

$$A \subseteq B \Leftrightarrow \mu_A(x) \leq \mu_B(x) \quad (3.3)$$

Nas operações com conjuntos fuzzy utilizam-se os conceitos de norma triangular (norma-t) e co-norma triangular (co-norma-t ou norma-s) (JANG, 1997). Uma norma-t é uma função $\tau: [0, 1] \times [0, 1] \rightarrow [0, 1]$ tal que, $\forall x, y, z, k \in [0, 1]$:

- 1) $(x \tau k) \leq (y \tau z)$, se $x \leq y, k \leq z$
- 2) $x \tau y = y \tau x$
- 3) $(x \tau y) \tau z = x \tau (y \tau z)$
- 4) $x \tau 0 = 0; x \tau 1 = x$

Uma co-norma-t, $\zeta: [0, 1] \times [0, 1] \rightarrow [0, 1]$, satisfaz (1) a (3) acima e, ainda:

$$5) x \zeta 0 = x; x \zeta 1 = 1$$

Exemplos de normas-t incluem o mínimo (*min* ou \wedge) e o produto algébrico (\bullet). Como exemplos de co-normas-t tem-se o máximo (*max* ou \vee) e a soma limitada (\oplus).

As normas e co-normas triangulares são utilizadas para representar a união e intersecção de conjuntos fuzzy. Utilizando-se os operadores *min* e *max* como norma-t e co-norma-t, respectivamente, tem-se:

União Fuzzy

A união de dois conjuntos fuzzy *A* e *B* é um conjunto fuzzy *C*, que pode ser escrito como $C = A \cup B$ cuja MF é relacionada aos conjuntos *A* e *B* por:

$$\mu_C(x) = \max(\mu_A(x), \mu_B(x)) = \mu_A(x) \vee \mu_B(x), \quad \forall x \in \mathcal{U} \quad (3.4)$$

Uma definição mais intuitiva, porém equivalente, da união fuzzy é o "menor" subconjunto fuzzy contendo *A* e *B* simultaneamente.

Intersecção Fuzzy

A intersecção de dois conjuntos fuzzy *A* e *B* é um conjunto fuzzy *C*, que pode ser escrito como $C = A \cap B$ cuja MF é relacionada aos conjuntos *A* e *B* por:

$$\mu_C(x) = \min(\mu_A(x), \mu_B(x)) = \mu_A(x) \wedge \mu_B(x) \quad \forall x \in \mathcal{U} \quad (3.5)$$

Da mesma forma que no caso da união, a interseção também pode ser definida de forma intuitiva como sendo o "maior" subconjunto fuzzy contendo A e B ao mesmo tempo.

Complemento Fuzzy

O complemento de um conjunto fuzzy A , denotado por \bar{A} é definido como:

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \quad \forall x \in \mathcal{U} \quad (3.6)$$

A Figura 3.1 (extraída de (JANG, 1997)) ilustra estas três operações básicas, onde na Figura 3.1(a) são mostrados os conjuntos fuzzy A e B ; na Figura 3.1(b) o conjunto complemento de A ; na Figura 3.1(c) a união de A e B ; e na Figura 3.1(d) a interseção de A e B .

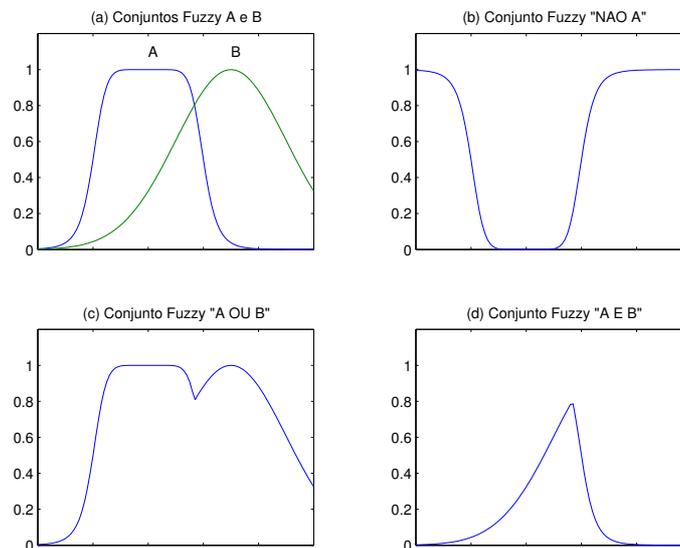


Figura 3.1: Operações com conjuntos fuzzy

Propriedades Algébricas

Utilizando as definições de união e interseção, implementadas pelos operadores max e min , e complemento, é possível verificar que as propriedades algébricas de involução, idempotência,

comutatividade, associatividade, distributividade e lei transitiva referentes a conjuntos ordinários também se aplicam a conjuntos fuzzy.

Considerando que as funções de pertinência dos conjuntos vazio e universo são definidas como sendo 0 e 1, respectivamente, $\forall x \in \mathcal{U}$, as seguintes propriedades também são válidas:

$$\left. \begin{array}{l} A \cap \emptyset = \emptyset \text{ e } A \cap \mathcal{U} = A \\ A \cup \emptyset = A \text{ e } A \cup \mathcal{U} = \mathcal{U} \end{array} \right\}$$

Ao contrário do observado em conjuntos clássicos, para conjuntos fuzzy tem-se que

$$A \cap \bar{A} \neq \emptyset \text{ e } A \cap \bar{A} \neq \mathcal{U}$$

caso \cap for tomada como sendo a interseção padrão.

3.1.3 Variáveis Lingüísticas

Uma variável lingüística é uma variável cujos valores são nomes de conjuntos fuzzy. Por exemplo, a temperatura de um determinado processo poderia ser uma variável lingüística 'Temperatura' assumindo valores *baixa*, *média*, *alta*, etc. Estes valores são descritos por de conjuntos fuzzy, conforme mostrado na Figura 3.2.

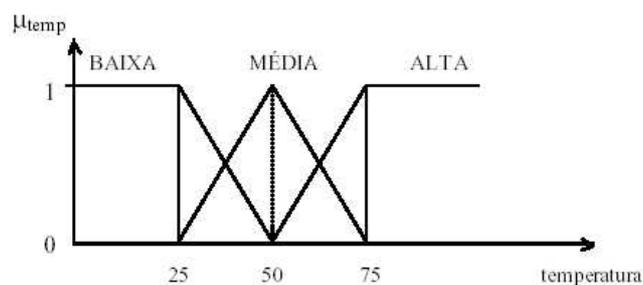


Figura 3.2: Variável lingüística *Temperatura*

Generalizando, os valores de uma variável lingüística podem ser sentenças em uma linguagem especificada, construídas a partir de termos primários (*alto*, *baixo*, *pequeno*, *médio*, *grande*, *zero*, por exemplo), de conectivos lógicos (negação *NÃO*, conectivos *E* e *OU*, conectivos mascarados, como *mas*, *porém*), de modificadores (como *muito*, *pouco*, *levemente*,

extremamente) e de delimitadores (como parênteses).

3.1.4 Proposições Fuzzy

Uma frase da forma (Ω é A), onde Ω é o nome de uma variável lingüística e A é um conjunto fuzzy definido no universo de discurso \mathcal{U} de Ω , é chamada de proposição fuzzy unária. No caso mais geral de uma proposição fuzzy n-ária, a representação se dá através do produto cartesiano das variáveis lingüísticas e da utilização de relações fuzzy ao invés de conjuntos fuzzy.

Proposições fuzzy podem ser combinadas por meio de diferentes operadores, como por exemplo, os conectivos lógicos E e OU , e o operador de implicação $SE \dots ENTÃO$; as proposições fuzzy daí resultantes podem ser descritas em termos de relações fuzzy. O valor de uma relação fuzzy, em função dos conjuntos fuzzy de cada operando, pode ser determinado de inúmeras maneiras.

Considerem-se as variáveis lingüísticas de nomes x e y , com universos de discurso X e Y , respectivamente, conjuntos fuzzy A e B definidos respectivamente em X e Y , e as proposições fuzzy (x é A) e (y é B). Conectando-se essas proposições através do operador OU , tem-se (x é A) OU (y é B), que pode ser escrito ((x,y) é $R_{A \text{ OUB}}$), onde $R_{A \text{ OUB}}$ é uma relação fuzzy em $X \times Y$, determinada por uma função de pertinência f_{OU} (usualmente uma co-norma-t):

$$R_{A \text{ OUB}} = \{ \mu_R(x,y) / (x,y) \} \text{ tal que } \mu_R(x,y) = f_{OU}(\mu_A(x), \mu_B(y)) \quad (3.7)$$

Conectando-se as proposições através do operador E , tem-se ((x,y) é $R_{A \text{ E } B}$), onde:

$$R_{A \text{ E } B} = \{ \mu_R(x,y) / (x,y) \} \text{ tal que } \mu_R(x,y) = f_E(\mu_A(x), \mu_B(y)) \quad (3.8)$$

onde f_E usualmente é uma norma-t.

O operador $SE \dots ENTÃO$ é também conhecido como declaração condicional e pode descrever a dependência do valor de uma variável lingüística em relação ao valor de outra. Em muitas aplicações as declarações condicionais são simplesmente denominadas de regras (Lingüísticas) e são, normalmente, frases da forma $SE (x \text{ é } A) ENTÃO (y \text{ é } B)$.

Esta frase, pode ser reduzida a uma proposição fuzzy binária ((x,y) é $R_{A \rightarrow B}$), na qual:

$$R_{A \rightarrow B} = \{\mu_R(x,y) / (x,y)\} \text{ tal que } \mu_R(x,y) = f_{\rightarrow}(\mu_A(x), \mu_B(y)) \quad (3.9)$$

onde f_{\rightarrow} é uma função de implicação.

Quando uma declaração condicional apresenta mais do que uma variável antecedente (x é A), as diversas variáveis são combinadas através do conectivo E , na forma:

$$SE ((x_1 \text{ é } A_1) E (x_2 \text{ é } A_2) E \dots E (x_m \text{ é } A_m)) ENTÃO (y \text{ é } B),$$

Várias declarações podem ser combinadas através dos conectivos OU e E, dependendo do operador de implicação:

$$\begin{aligned} &SE ((x \text{ é } A_1) ENTÃO (y \text{ é } B_1)) \\ &OU \\ &SE ((x \text{ é } A_2) ENTÃO (y \text{ é } B_2)) \\ &OU \\ &\vdots \\ &SE ((x \text{ é } A_n) ENTÃO (y \text{ é } B_n)) \end{aligned}$$

3.2 Sistemas de Inferência Fuzzy

O **Sistema de Inferência Fuzzy** é uma estrutura computacional baseada nos conceitos de teoria de conjuntos fuzzy, regras fuzzy $SE \dots ENTÃO$ e raciocínio fuzzy (JANG, 1997). Este sistemas tem sido aplicados em uma grande variedade de campos, tais como controle automático, classificação de dados, sistemas especialistas, predição de séries temporais, robótica e reconhecimento de padrões.

A estrutura básica de um sistema de inferência fuzzy consiste de três camadas conceituais:

1. Uma base de regras, que contém um conjunto de regras fuzzy;
2. Uma base de dados (ou dicionário), que define as funções de pertinência usadas nas regras fuzzy;
3. Um mecanismo de raciocínio, que desenvolve o procedimento de inferência com base nas regras e fatos conhecidos, possibilitando uma saída.

As entradas do sistema de inferência fuzzy podem ser tanto valores fuzzy (ver Seção 3.1) como valores crisp (que são tratados como *singleton* fuzzy), porém as saídas produzidas são

quase sempre um conjunto fuzzy. Em algumas situações é necessário que a saída seja um valor real. Para isso torna-se necessário a aplicação de um método de desfuzificação para obter um valor crisp que melhor represente o conjunto fuzzy. Na Figura 3.3 , extraída de JANG, (1997), é apresentado um sistema de inferência fuzzy com saída crisp.

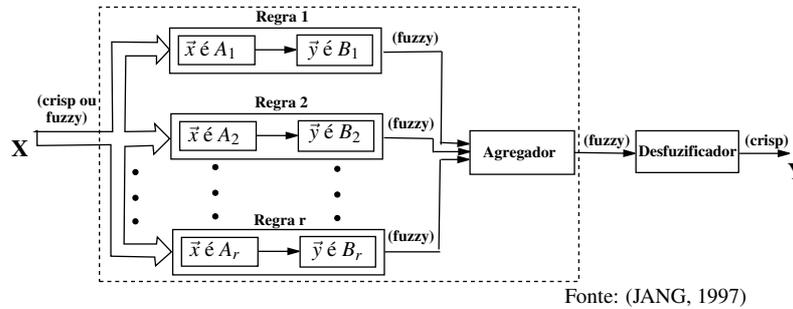


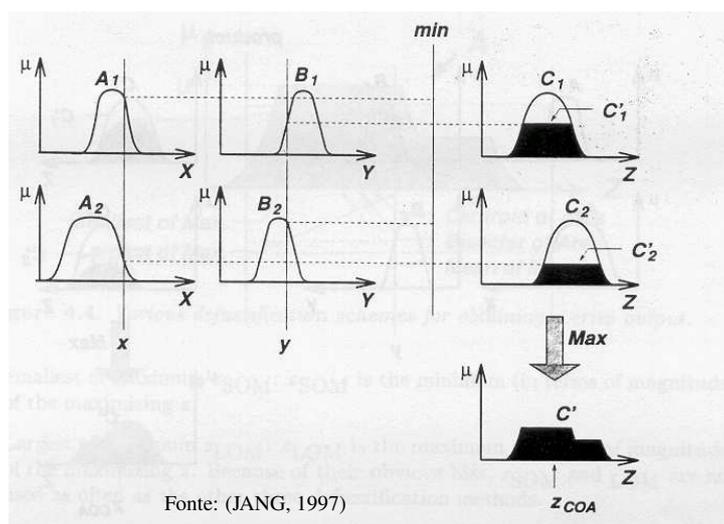
Figura 3.3: Diagrama de Blocos de um Sistema de Inferência Fuzzy

Um sistema de inferência fuzzy implementa um mapeamento não linear entre seu espaço de entrada e o espaço de saída. Este mapeamento é composto por um número de regras fuzzy *SE ... ENTÃO*, onde cada uma delas descreve o comportamento local do mapeamento. Em particular, o antecedente de uma regra define uma região no espaço de entrada, enquanto que o conseqüente especifica a saída no espaço de saída.

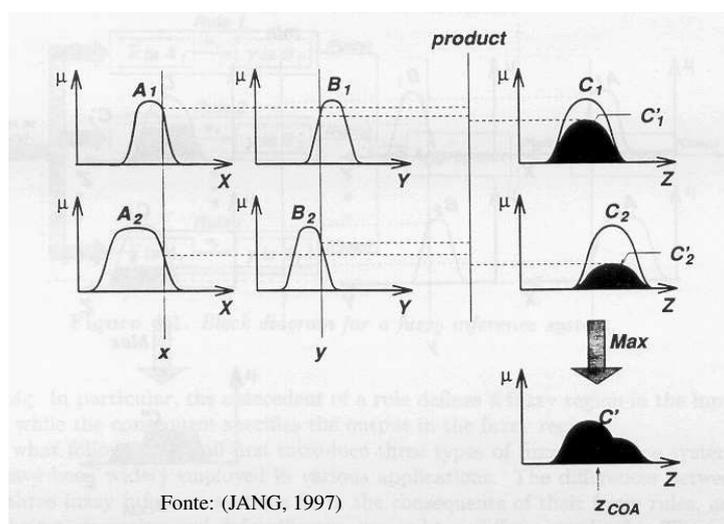
A seguir são apresentados três tipos usuais de sistemas de inferência fuzzy que são empregados em uma larga classe de aplicações.

3.2.1 Modelo Fuzzy de Mamdani

O sistema de Inferência Fuzzy de Mamdani (MAMDANI 1973) foi proposto como o primeiro sistema usado na tentativa de controlar uma máquina a vapor e caldeira por meio de um conjunto de regras linguistics de controle, obtidas a partir da experiência de operadores humanos. A Figura 3.4 ilustra como um sistema de inferência de Mamdani de duas regras infere uma saída final z a partir de duas entradas x e y . Na Figura 3.4(a) a inferência é feita utilizando-se \min e \max para operadores norma-t e co-norma-t, respectivamente, enquanto que na Figura 3.4(b) a inferência é feita usando *produto* e \max para operadores norma-t e co-norma-t, respectivamente.



(a)



(b)

Figura 3.4: Sistemas de inferência fuzzy de Mamdani

Desfuzificação

Desfuzificação é o mecanismo pelo qual um valor real é extraído a partir de um conjunto fuzzy. Existem cinco métodos para desfuzificação de um conjunto fuzzy A pertencente a um universo de discurso Z , conforme mostrado na Figura 3.5, onde o conjunto A resulta de uma combinação das funções de pertinência de saída, tal como C' nas Figuras 3.4(a) e 3.4(b).

Este cinco métodos de desfuzificação são apresentados a seguir e podem ser visualizados na Figura 3.5:

Centro de área (z_{COA}):

$$z_{COA} = \frac{\int_z \mu_A(z)zdz}{\int_z \mu_A(z)dz} \quad (3.10)$$

onde $\mu_A(z)$ é a função de pertinência da saída agregada. O centróide é amplamente adotado como estratégia de desfuzificação, o qual é remanescente do cálculo do valor esperado da distribuição probabilística.

Bissetor de área (z_{BOA}): O valor de z_{BOA} satisfaz:

$$\int_{\alpha}^{z_{BOA}} \mu_A(z)dz = \int_{z_{BOA}}^{\beta} \mu_A(z)dz \quad (3.11)$$

onde $\alpha = \min\{z | z \in Z\}$ e $\beta = \max\{z | z \in Z\}$. Isto é, a linha vertical $z = z_{BOA}$ divide as regiões entre $z = \alpha$, $z = \beta$, $y = 0$ e $y = \mu_A(z)$ em duas regiões com a mesma área.

Média do Máximo (z_{MOM}): O valor de z_{MOM} é a média do máximo de z em que a função de pertinência alcança o máximo μ^* , ou seja:

$$z_{MOM} = \frac{\int_{Z'} zdz}{\int_{Z'} dz}, \quad (3.12)$$

onde $Z' = \{z | \mu_A(z) = \mu^*\}$. Em particular, se $\mu_A(z)$ tem um valor máximo único em $z = z^*$, então $z_{MOM} = z^*$. Além disso, se $\mu_A(z)$ alcança seu máximo sempre que $z \in [z_{esquerda}, z_{direita}]$, então $z_{MOM} = (z_{esquerda} + z_{direita}) / 2$. A média dos máximos é a estratégia de desfuzificação utilizada no controlador fuzzy de Mamdani.

Menor do Máximo (z_{SOM}): O valor de z_{SOM} é o mínimo de z onde ocorre o máximo $\mu_A(z) = \mu^*$.

Maior do Máximo (z_{LOM}): O valor de z_{LOM} é o máximo de z onde ocorre o máximo $\mu_A(z) = \mu^*$.

Devido à suas limitações, z_{SOM} e z_{LOM} não são usados tão freqüentemente quanto os três outros métodos de desfuzificação.

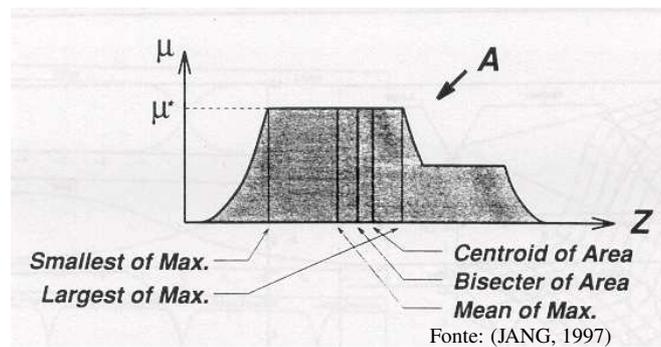


Figura 3.5: Métodos de desfuzificação

3.2.2 Modelo Fuzzy de Takagi e Sugeno

O modelo fuzzy Takagi e Sugeno (também conhecido como o modelo TS), foi proposto por Takagi e Sugeno (TAKAGI e SUGENO, 1985) num esforço de desenvolver uma abordagem sistemática para gerar regras fuzzy a partir de um dado conjunto de dados de entrada e saída. Uma regra típica num modelo TS tem a forma:

$$SE \ x \ \acute{e} \ A \ E \ y \ \acute{e} \ B \ \text{ ENT\AAO } \ z = f(x,y),$$

onde A e B são conjuntos fuzzy enquanto que $z = f(x,y)$ é uma função. Geralmente $f(x,y)$ é um polinômio nas variáveis de entrada x e y , mas pode ser qualquer função que descreva a saída do modelo dentro da região fuzzy especificada pelos antecedentes da regra. Quando $f(x,y)$ é um polinômio de primeira ordem, o sistema de inferência resultante é chamado modelo fuzzy de Sugeno de primeira ordem, que foi originalmente proposto. Quando f é uma constante, temos um modelo fuzzy de Sugeno de ordem zero, que pode ser visto como um caso especial do sistema de inferência fuzzy de Mamdani, em que cada conseqüente da regra é especificado por um *singleton* fuzzy (ou um conseqüente pré-desfuzificado), ou um caso especial do modelo fuzzy Tsukamoto, em que cada conseqüente da regra é especificado por uma função de pertinência do tipo degrau no centro constante. Além disso, um modelo fuzzy de Sugeno de ordem zero é, sob certas restrições fracas, funcionalmente equivalente a uma rede de função base radial (ver Seção 3.5).

A Figura 3.6 mostra o procedimento do raciocínio fuzzy para um modelo fuzzy de Sugeno de primeira ordem. Uma vez que cada regra tem uma saída crisp, a saída total é obtida via média ponderada, assim evitando o consumo de tempo da desfuzificação requerida num

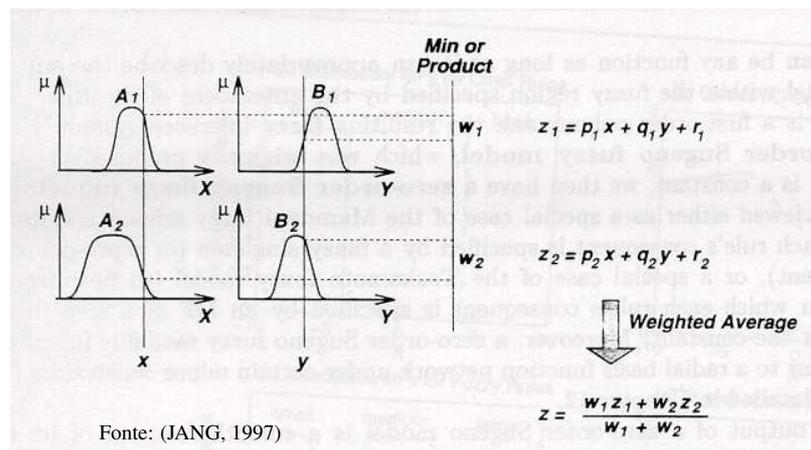


Figura 3.6: Sistemas de inferência fuzzy de Takagi e Sugeno de Primeira Ordem

modelo fuzzy de Mamdani. Na prática, o operador média ponderada é algumas vezes trocado pelo operador soma ponderada (isto é, $z = w_1 z_1 + w_2 z_2$ na Figura 3.6) para reduzir cálculos futuros, especialmente no treinamento de um sistema de inferência.

Devido a não necessidade do processo de desfuzificação, o modelo fuzzy de Sugeno é um forte candidato para modelos fuzzy baseados em regras.

3.2.3 Modelo Fuzzy de Tsukamoto

No modelo fuzzy de Tsukamoto (TSUKAMOTO, 1979), o conseqüente de cada regra fuzzy *SE ... ENTÃO* é um conjunto fuzzy com função de pertinência monotônica, como mostrado na Figura 3.7. Como resultado, a saída inferida de cada regra é definida como um valor crisp induzida pela intensidade de disparo da regra. A saída total é tomada como a média ponderada das saídas das regras. A Figura 3.7 ilustra o procedimento do raciocínio para um sistema de duas regras e duas entradas. Desde que cada regra infere uma saída crisp, o modelo fuzzy de Tsukamoto agrega cada saída das regras pelo método da média ponderada e assim evita consumir tempo no processo de desfuzificação. Contudo, o modelo fuzzy de Tsukamoto não é usado freqüentemente, uma vez que ele não é tão transparente como os modelos fuzzy de Mamdani a Sugeno.

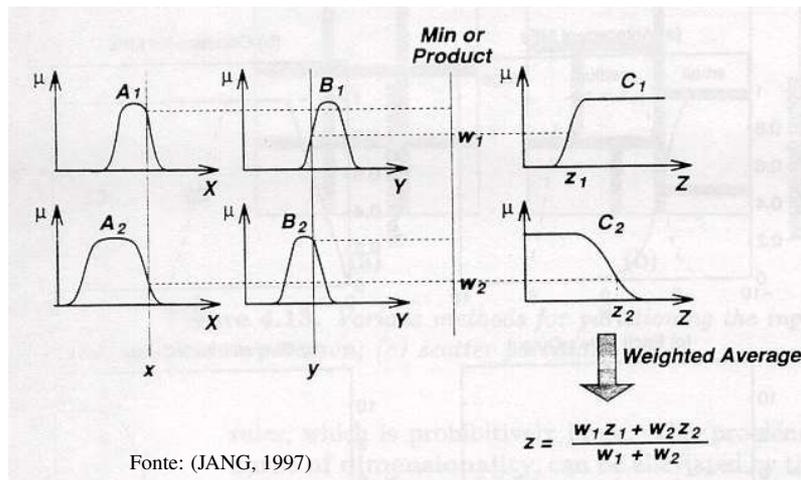


Figura 3.7: Sistemas de Inferência Fuzzy de Tsukamoto

3.3 Sistema de Inferência Fuzzy Neuro-Adaptativo

O sistema de inferência fuzzy neuro-adaptativo (ANFIS²) foi proposto por J.-S. R. Jang em 1993 (JANG, 1993) com a finalidade de criar uma estrutura adaptável para o sistema de inferência fuzzy tradicional, implementando um sistema fuzzy Sugeno de primeira ordem (TAKAGI e SUGENO, 1985). A necessidade deste tipo de estrutura se baseia em alguns aspectos da abordagem fuzzy convencional, tais como:

- Não existe método padrão para transformar a experiência ou o conhecimento humano em uma base de regras a serem utilizadas em sistema de inferência fuzzy;
- Existe uma grande necessidade de se criar métodos que ajustem as funções de pertinência (MF's) com a finalidade de minimizar o erro de saída ou aumentar índices de performance.

Neste sentido, o objetivo dos sistemas fuzzy neuro-adaptáveis é servir de base para a construção de um sistema de inferência fuzzy que apresente uma melhor performance quando aplicado a um problema específico, sendo ajustado a este problema a partir de pares entrada-saída pré-determinados.

² do inglês *Adaptive-Neuro-Based Fuzzy Inference System*

3.3.1 Arquitetura ANFIS

Como citado anteriormente, o ANFIS implementa um modelo fuzzy Sugeno de primeira ordem. Por questões de simplicidade, será considerado um sistema fuzzy com duas entradas x e y , e uma saída z . Para um modelo fuzzy Sugeno de primeira ordem a base de regras com duas regras *SE ... ENTÃO* é construída da seguinte forma:

$$\begin{aligned} \text{Regra1: } & \text{SE } x \text{ é } A_1 \text{ E } y \text{ é } B_1 \text{ ENTÃO } f_1 = p_1x + q_1y + r_1 \\ \text{Regra2: } & \text{SE } x \text{ é } A_2 \text{ E } y \text{ é } B_2 \text{ ENTÃO } f_2 = p_2x + q_2y + r_2 \end{aligned}$$

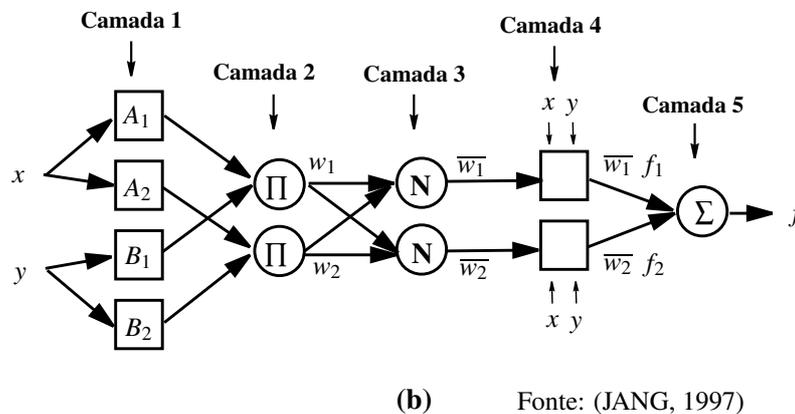
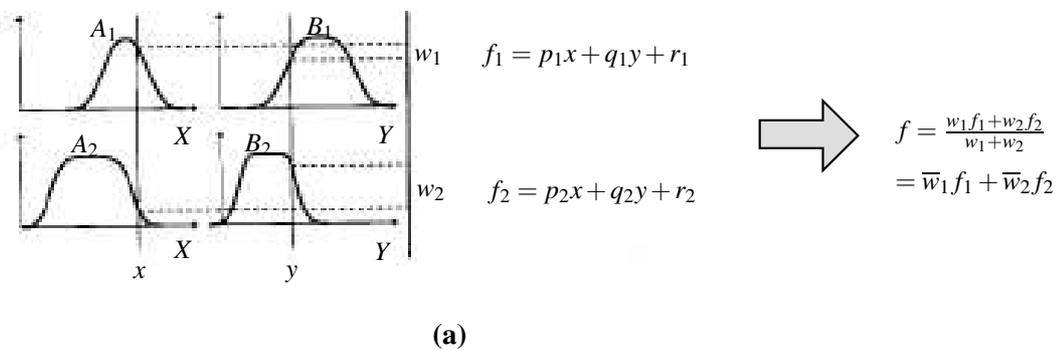


Figura 3.8: (a) Modelo fuzzy Sugeno de primeira ordem de duas entradas com duas regras; (b) Estrutura ANFIS correspondente

A Figura 3.8(a) ilustra o mecanismo de raciocínio para este modelo fuzzy Sugeno. A estrutura ANFIS equivalente a este modelo é mostrada na Figura 3.8(b), onde os nós de uma mesma camada possuem funções similares, conforme descrito a seguir:

Camada 1: Cada nó i desta camada (denotado por um quadrado na Figura 3.8) representa

um subconjunto fuzzy da variável lingüística de entrada (A_1 e A_2 para X e B_1 e B_2 para Y) e fornece o grau de pertinência de uma dada entrada (x ou y) no i -ésimo subconjunto fuzzy, ou seja:

$$\begin{aligned} O_{1,i} &= \mu_{A_i}(x), & \text{para } i = 1, 2 \text{ ou} \\ O_{1,i} &= \mu_{B_{i-2}}(y), & \text{para } i = 3, 4 \end{aligned} \quad (3.13)$$

onde $O_{\ell,i}$, com $\ell = 1, \dots, 5$ e $i = 1, \dots, 4$, representa a saída do i -ésimo nó da ℓ -ésima camada, x (ou y) é a entrada para o nó i , A_i (ou B_{i-2}) representa o subconjunto fuzzy (tal como "pequeno", "grande") associado a este nó e $\mu_{A_i}(x)$ (ou $\mu_{B_{i-2}}(y)$) é a função de pertinência do subconjunto i . A função de pertinência deve ser escolhida de forma a melhor se ajustar ao problema a ser resolvido pelo sistema. No exemplo mostrado na Figura 3.8a função de pertinência mostrada é do tipo sino, e pode ser obtida por:

$$\mu_A(x) = \frac{1}{1 + \left| \frac{x-c_i}{a_i} \right|^{2b_i}}, \quad (3.14)$$

onde $\{a_i, b_i, c_i\}$ é o conjunto de parâmetros desta função. Os parâmetros nesta camada são chamados **parâmetros premissa**.

Camada 2: Cada nó i desta camada (denotado por um círculo rotulado por Π) tem como saída o produto de todos os sinais que entram nesta camada, isto é:

$$O_{2,i} = w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), \quad i = 1, 2. \quad (3.15)$$

Estes nós representam o grau de ativação da i -ésima regra.

Camada 3: Cada nó i desta camada (denotado por um círculo rotulado por N) calcula a relação entre o i -ésimo grau de ativação da regra e o somatório de todos os graus de ativação, da forma:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad (3.16)$$

As saídas desta camada são chamadas de **graus de ativação normalizados**.

Camada 4: Cada nó i nesta camada (representado por um quadrado na Figura 3.8) executa a seguinte função;

$$O_{i,4} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + \tau_i) \quad (3.17)$$

onde \bar{w}_i é o i -ésimo grau de ativação normalizado, calculado na camada 3, e $\{p_i, q_i, \tau_i\}$ é o conjunto de parâmetros desta camada. Estes parâmetros são normalmente referidos como **parâmetros consegüentes**.

Camada 5: O único nó (denotado por um círculo rotulado por Σ) que forma esta camada calcula a saída geral como sendo o somatório de todos os sinais que entram nesta camada, isto é:

$$O_{5,1} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (3.18)$$

Desta forma é possível construir uma rede adaptativa que é funcionalmente equivalente a um sistema de inferência fuzzy Sugeno de primeira ordem.

Para desempenhar o ajuste desse sistema é necessário aplicar um algoritmo de aprendizado híbrido. Esta técnica fornece um método para que o modelo fuzzy "aprenda" informações a respeito de uma base de dados, com a finalidade de calcular os parâmetros das funções de pertinência fuzzy que melhor se ajustam à um conjunto de dados entrada/saída. Este tipo de treinamento funciona similarmente aos das redes neurais, porém combinando algoritmos em um treinamento híbrido, conforme apresentado em (JANG, 1997).

3.4 Redes Neurais de Múltiplas Camadas

Redes Neurais de Múltiplas Camadas são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência.

Em uma rede neural de múltiplas camadas são identificadas a camada de entrada, as camadas intermediárias (ou camadas ocultas) e a camada de saída, como mostrado na Figura 3.9.

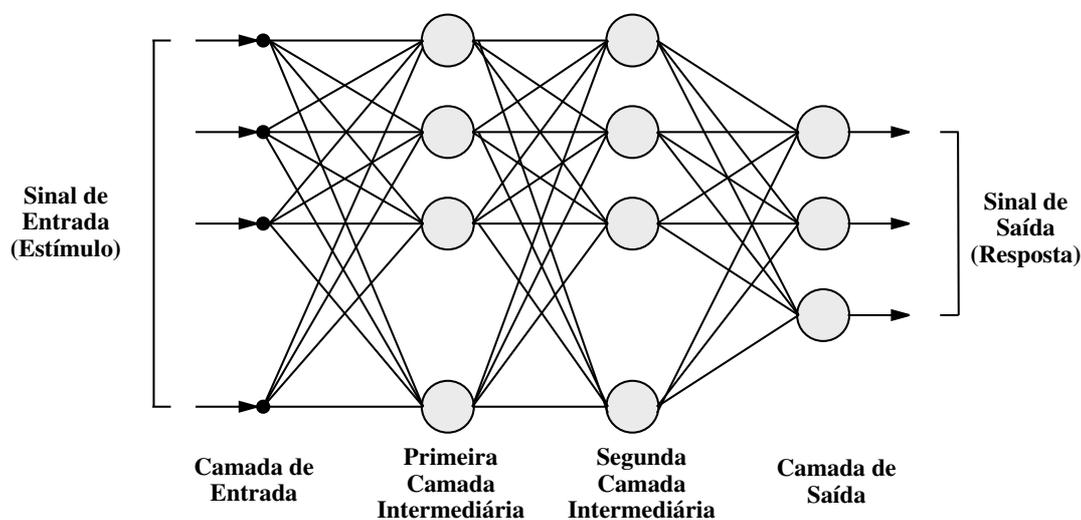


Figura 3.9: Diagrama ilustrativo de uma Rede Neural com duas camadas intermediárias

Neste tipo de rede um sinal de entrada se propaga para frente através da rede, camada por camada, até que a camada de saída seja atingida, gerando assim a resposta da rede a este sinal de entrada. Este processo é também conhecido como *feedforward*.

As redes neurais de múltiplas camadas, normalmente chamadas de *perceptrons de múltiplas camadas* (MLP's³), tem sido aplicadas com sucesso para resolver diversos tipos de problemas de difícil solução, através do seu treinamento de forma supervisionada, que utiliza um algoritmo conhecido como *Algoritmo Backpropagation*, apresentado a seguir.

3.4.1 O Algoritmo *Backpropagation*

Durante o treinamento com o algoritmo *backpropagation*, a rede opera em uma seqüência de dois passos. Primeiro, um padrão é apresentado à camada de entrada da rede. A atividade resultante flui através da rede, camada por camada, até que a resposta seja produzida pela camada de saída. No segundo passo, a saída obtida é comparada à saída desejada para esse padrão particular. Se esta não estiver correta, o erro é calculado e propagado a partir da camada de saída até a camada de entrada. Desta forma, os pesos das conexões das unidades

³ do inglês *Multilayer Perceptrons*

das camadas internas vão sendo modificados conforme o erro é retropropagado.

Este processo gira em torno de um conjunto de amostras de treinamento $\{x(p), d(p)\}$ (para $n = 1, \dots, P$), onde P é o número de amostras de treinamento contidas no conjunto de amostras, e é baseado no método do gradiente descendente (HAYKIN, 2001), e pode ser resumido como segue:

Primeiro Passo \implies Inicialização: Caso não exista nenhuma informação prévia armazenada na rede, é necessário que os pesos sinápticos e limiares sejam iniciados a partir de uma distribuição uniforme dentro de um determinado intervalo pré-estabelecido;

Segundo Passo \implies Apresentação dos Exemplos de treinamento: Após a inicialização um padrão de treinamento $\{x(p), d(p)\}$, onde $x(p)$ é o vetor de entrada da rede e $d(p)$ é a resposta que se deseja obter para o estímulo $x(p)$, deve ser apresentado à rede;

Terceiro Passo \implies Propagação para Frente (*Feedforward*): Neste passo a saída da rede é calculada propagando-se o sinal camada por camada. Assim, a saída do neurônio j na camada ℓ , denotado por $y_j^{(\ell)}(p)$, pode ser expressa por:

$$y_j^{(\ell)}(p) = \varphi_j \left(v_j^{(\ell)}(p) \right) \quad (3.19)$$

onde

$$v_j^{(\ell)}(p) = \sum_{i=0}^m \omega_{ji}^{(\ell)}(p) y_i^{(\ell-1)}(p)$$

e m é o número total de entradas incluindo os limiares, $\omega_{ji}^{(\ell)}(p)$ é o peso que parte da unidade i da camada $\ell - 1$ para a unidade j da camada ℓ , $y_i^{(\ell-1)}(p)$ é a saída do neurônio i na camada anterior $\ell - 1$, para o p -ésimo padrão de treinamento e $\varphi_j(\cdot)$ é a função de ativação associada ao neurônio j . Para $i = 0$, temos $y_0^{(\ell-1)}(p) = +1$ e $\omega_{j0}^\ell = b_j^{(\ell)}(p)$ que é o limiar aplicado ao neurônio j na camada ℓ ,

Se o neurônio j está na primeira camada intermediária, ou seja $\ell = 1$, então:

$$y_j^{(0)}(p) = x_j(p) \quad (3.20)$$

onde $x_j(p)$ é o j -ésimo elemento do vetor de entrada $x(p)$. Se o neurônio j está na camada de saída, isto é $\ell = L$ (onde L é o número de camadas da rede), então:

$$y_j^{(L)} = o_j(p) \quad (3.21)$$

onde $o_j(p)$ é o j -ésimo elemento do vetor de saída, obtido pela propagação do vetor de entrada $x(p)$ através da rede.

Desta forma é possível calcular o sinal de erro da saída $e_j(p)$

$$e_j(p) = d_j(p) - o_j(p), \quad (3.22)$$

onde $d_j(p)$ é o j -ésimo elemento do vetor de saídas desejadas $d(p)$.

Quarto Passo \implies Propagação para trás (*Backpropagation*): Neste passo os gradientes locais da rede δ são calculados por

$$\delta_j^\ell(p) = \begin{cases} e_j(p)\varphi'_j\left(v_j^{(L)}(p)\right) & \text{para o neurônio } j \text{ da camada de saída } L \\ \varphi'\left(v_j^{(\ell)}(p)\right) \sum_k \delta_k^{(\ell+1)}(p)\omega_{kj}^{(\ell+1)}(p) & \text{para o neurônio } j \text{ na camada intermediária } \ell \end{cases} \quad (3.23)$$

onde o apóstrofe em $\varphi'_j(\circ)$ representa a diferenciação em relação ao argumento. O ajuste dos pesos sinápticos da rede na camada ℓ é feito de acordo com a regra delta generalizada, como segue:

$$\omega_{ji}^{(\ell)}(p)(k+1) = \omega_{ji}^{(\ell)}(p)(k) + \Delta\omega_{ij}(p)(k) \quad (3.24)$$

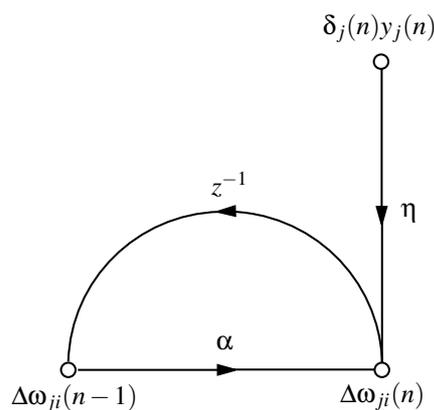
onde

$$\Delta\omega_{ji}(p)(k) = \eta\delta_j^{(\ell)}(p)(k)y_i^{(\ell-1)}(p)(k) \quad (3.25)$$

e η é o parâmetro *taxa de aprendizagem*, que representa a quantidade de modificação dos pesos na iteração k . Quanto menor for o parâmetro da taxa de aprendizagem, menor serão as variações dos pesos sinápticos da rede, de uma iteração para outra, e mais suave será a trajetória no espaço de pesos. Esta melhoria, entretanto, é obtida à custa de uma aprendizagem lenta. Por outro lado, se fizermos o parâmetro da taxa de aprendizagem muito grande, para acelerar o processo de aprendizagem, as grandes modificações nos pesos sinápticos resultantes podem tornar a rede instável (i.e., oscilatória). Um método simples de aumentar a taxa de aprendizagem, evitando no entanto o perigo de instabilidade é modificar a Equação (3.25) para seguinte forma:

$$\Delta\omega_{ji}(p)(k) = \alpha \left[\Delta\omega_{ji}^{(\ell)}(p)(k-1) \right] + \eta\delta_j^{(\ell)}(p)(k)y_i^{(\ell-1)}(p)(k) \quad (3.26)$$

onde α é usualmente um número positivo chamado de *constante de momento*. A constante de momento controla o laço realimentação que age em torno de $\omega_{ji}(n)$ (HAYKIN, 2001), como mostrado na Figura 3.10, onde z^{-1} é o operador de atraso unitário.



Fonte: (HAYKIN, 2001)

Figura 3.10: Fluxo de sinal ilustrando o efeito da constante de momento

Quinto Passo \implies Iterações: As propagações para frente e para trás, (passos 3 e 4, respec-

tivamente) devem ser calculadas, apresentado-se novos padrões de treinamento para a rede, até que um critério de parada (tais como erro máximo admissível, número máximo de iterações, etc., ou a combinação deles) seja satisfeito.

3.5 Redes Neurais de Função de Base Radial

Uma Rede Neural de Base Radial (RBFN⁴) é uma classe de RNA's que baseia-se no modelo de campos de recepção sobrepostos, presentes em regiões do córtex cerebral ligadas a processamento de informações sensoriais (JANG, 1997, p.238). Seu rápido processo de aprendizagem e seu grande poder de generalização tem motivado o uso deste tipo de rede para aproximação funcional de sistemas dinâmicos não-lineares.

O princípio deste tipo de rede baseia-se no fato de que é possível encontrar uma superfície em um espaço multidimensional que produz o melhor ajuste para os dados de treinamento (HAYKIN, 1994, p.236). Sua arquitetura básica é ilustrada na Figura 3.11 e constitui-se de três camadas, sendo a primeira camada constituída por p entradas (x_j com $j = 1, \dots, p$); a segunda, por n neurônios ocultos R_i , associados a n vetores centros \mathbf{u}_i , com $i = 1, \dots, n$, e a terceira, uma camada constituída por k neurônios de saída (f_k com $k = 1, \dots, \ell$).

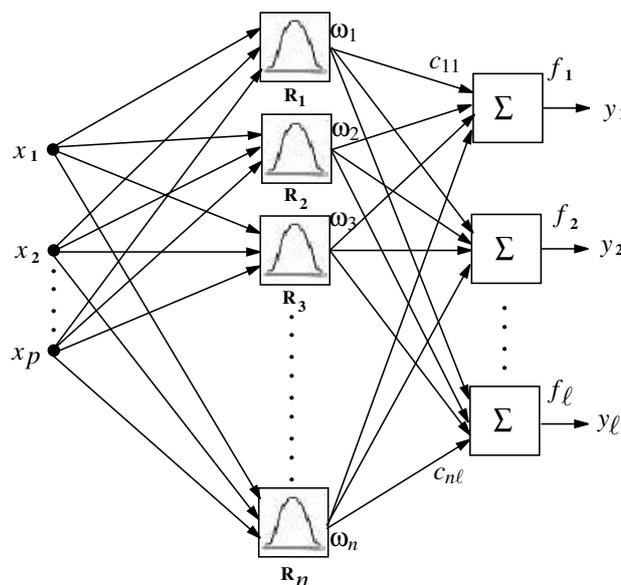


Figura 3.11: Estrutura Básica de uma RBFN

⁴ do inglês *Radial Basis Function Network*

Na figura 3.11, cada neurônio oculto R_i forma uma estrutura chamada "campo receptivo", conforme mostrado na Figura 3.12.

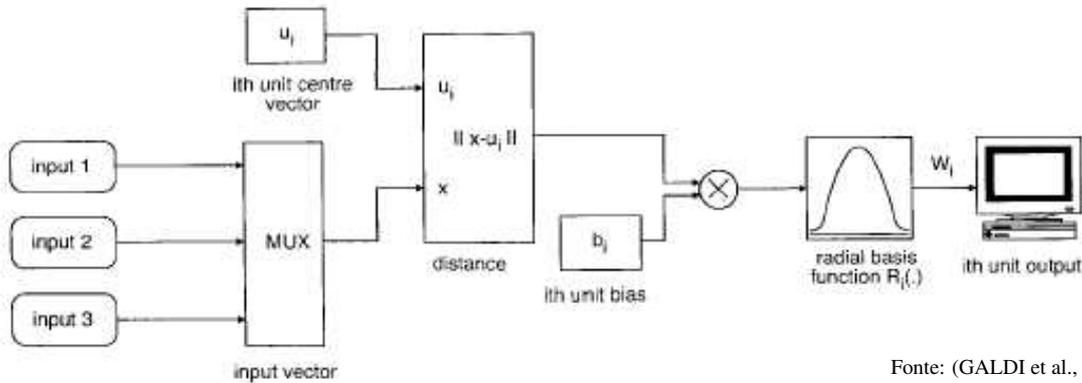


Figura 3.12: Modelo de Campo Receptivo

Em cada campo receptivo é calculada a distância entre a entrada (\mathbf{x}) e seu respectivo centro (\mathbf{u}_i), conforme figura 3.12. Após esse passo, é calculado o nível de ativação ω_i , que representa o grau de similaridade entre a entrada da rede e o centro \mathbf{u}_i . Segundo (JANG, 1997) este valor pode ser expresso da seguinte forma:

$$\omega_i = R_i(\mathbf{x}) = R_i(\|\mathbf{x} - \mathbf{u}_i\|/\sigma_i) \quad i = 1, 2, \dots, n \quad (3.27)$$

onde $R_i(\cdot)$ é a i -ésima função de base radial com valor máximo unitário na origem, como mostrado na Figura 3.13.

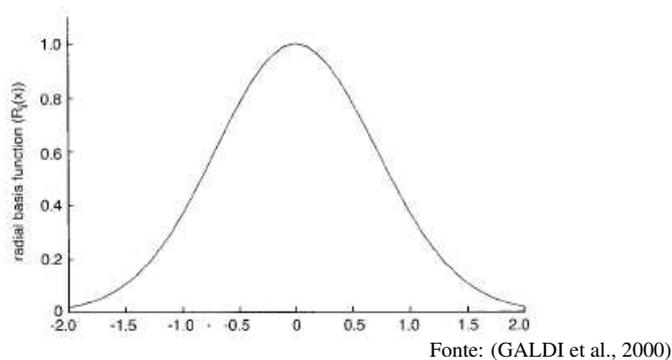


Figura 3.13: Exemplo de Função de Base Radial

Normalmente, a função $R_i(\cdot)$ é uma função Gaussiana

$$R_i(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{u}_i\|^2}{2\sigma_i^2}\right) \quad (3.28)$$

ou uma função logística

$$R_i(\mathbf{x}) = \frac{1}{1 + \exp(\|\mathbf{x} - \mathbf{u}_i\|^2/\sigma_i^2)} \quad (3.29)$$

Assim, o nível de ativação da função de base radial ω_i calculada para a i -ésima unidade oculta é máxima quando a entrada \mathbf{x} é igual ao centro \mathbf{u}_i .

Para se calcular a saída final da RBFN pode-se usar diferentes metodologias. Neste trabalho a saída final da RBFN será calculada como sendo a soma ponderada dos níveis de ativação de cada campo receptivo, sendo expressa da seguinte forma:

$$\begin{aligned} y_k &= \sum_{i=1}^n c_{ik} \omega_i = \sum_{i=1}^n c_{ik} R_i(\mathbf{x}) \\ &= \sum_{i=1}^n c_{ik} R_i(\|\mathbf{x} - \mathbf{u}_i\|/\sigma_i) \quad \text{para } k = 1, \dots, \ell. \end{aligned} \quad (3.30)$$

onde c_{ik} é o peso de conexão entre a saída do campo receptivo i e a unidade de saída k .

Diversos algoritmos de treinamento tem sido propostos para identificar os parâmetros (\mathbf{u}_i , σ_i e c_i) de uma rede de base radial. Neste trabalho será aplicado um procedimento de treinamento baseado no método ortogonal dos mínimos quadrados (CHEN et. al., 1991). Este procedimento ajusta os centros das funções de base radial um por um de uma forma racional, até que uma rede adequada seja obtida.

3.6 Mapas Auto Organizáveis

Segundo (HAYKIN, 2001), em um mapa auto-organizável, os neurônios estão colocados em nós de uma *grade* que normalmente é uni- ou bidimensional. Mapas de dimensionalidade mais alta são também possíveis, mas não são comuns. Os neurônios se tornam *seletivamente sintonizados* a padrões de entrada (estímulos) ou classes de padrões de entrada no decorrer de um processo de aprendizagem. Um mapa auto-organizável é, portanto, caracterizado pela

formação de um mapa topográfico dos padrões de entrada no qual as localizações espaciais (i.e., coordenadas) dos neurônios na grade são indicativas das características estatísticas intrínsecas contidas nos padrões de entrada (KOHONEN, 1991).

3.6.1 O Modelo de Kohonen

O modelo de Kohonen produz um mapeamento topológico que localiza otimamente um número fixo de vetores em um espaço de entrada de dimensionalidade mais elevada, e desse modo facilita a compressão de dados. Segundo (HAYKIN, 2001) é possível derivar o modelo de Kohonen de duas formas. A primeira utiliza as idéias básicas de auto-organização, motivadas por considerações neuro-biológicas, para derivar o modelo tradicional (KOHONEN, 1991). Figura 3.14 ilustra este modelo.

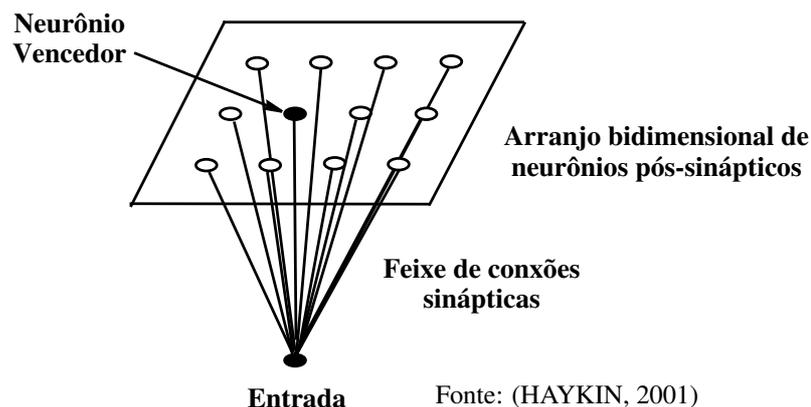


Figura 3.14: Diagrama Ilustrativo de Mapa de Kohonen

A segunda forma de se derivar o modelo de Kohonen é utilizando uma abordagem de quantização vetorial (LQV⁵) que usa um modelo envolvendo um codificador e um decodificador, e é motivada por considerações da teoria de comunicação (LUTTRELL, 1991). Na Figura 3.15 é mostrado o modelo LQV de Kohonen, também chamado *Camada de Kohonen*. Nesta seção será tratada mais especificamente esta segunda abordagem.

⁵ do inglês *Learning Vector Quantization*

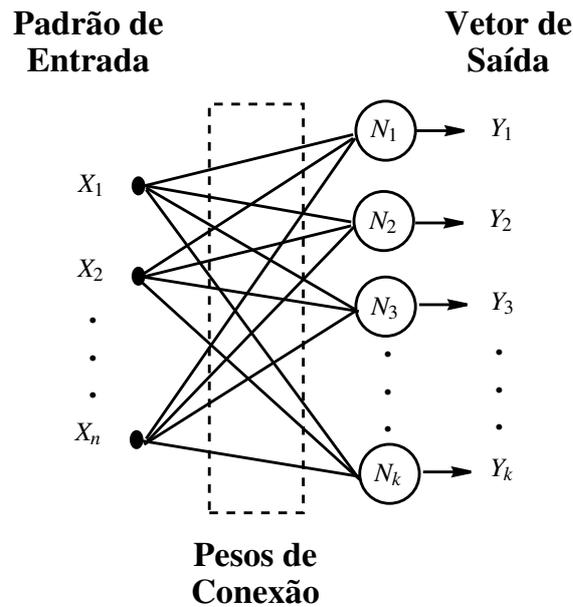


Figura 3.15: Diagrama Ilustrativo de Camada de Kohonen

A Camada de Kohonen

Na camada de Kohonen (KOHONEN, 1991), para um dado vetor de entrada " $x = [x_1, x_2, \dots, x_n]$ ", um e somente um neurônio de Kohonen é ativado em "1", os demais apresentam a saída zero, isto é, para um dado vetor de entrada existe um único neurônio ativo.

Todo neurônio de Kohonen está associado a um conjunto de pesos que o conectam a cada componente do vetor de entrada, como ilustrado na Figura 3.15. Por exemplo o neurônio N_1 , tem associado a ele seus respectivos pesos $(w_{11}, w_{21}, \dots, w_{n1})$, compreendendo um vetor peso $W_1 \in \mathbb{R}^n$. A saída de cada neurônio da camada de Kohonen é a soma das entradas, ponderadas pelos pesos de conexão, como expressado em 3.31.

$$Y_j = \sum_{i=1}^n X_i W_{ij} \quad (3.31)$$

onde Y_j é a saída do j -ésimo neurônio de Kohonen.

A rede de Kohonen organiza os vetores de entrada em grupos que são similares. Isto é acompanhado pelo ajuste dos pesos da rede de Kohonen, de modo que os vetores de entrada similares ativam o mesmo neurônio de Kohonen. O treinamento da rede é feito por um algoritmo não supervisionado, a isto se deve a dificuldade de determinar qual neurônio específico do camada de Kohonen será ativado, porém, somente é necessário garantir que o treinamento

separe os vetores em grupos. O neurônio a ser ativado é determinado por $\max(Y_j)$ e o ajuste dos pesos é feito através da Equação 3.32.

$$W_{ij}(k+1) = W_{ij}(k) + \alpha(k)(X_i - W_{ij}(k)) \quad , \quad 0 < \alpha \leq 1 \quad (3.32)$$

Inicialmente toma-se α com valor próximo do limite superior no intervalo $[0,1]$, e a medida que os pesos vão se ajustando o valor de α é reduzido para um valor próximo de "0".

Como as entradas da rede de kohonen são valores normalizados, este processo pode ser entendido como uma aproximação de vetores contidos em um círculo de raio unitário, conforme mostrado na Figura 3.16.

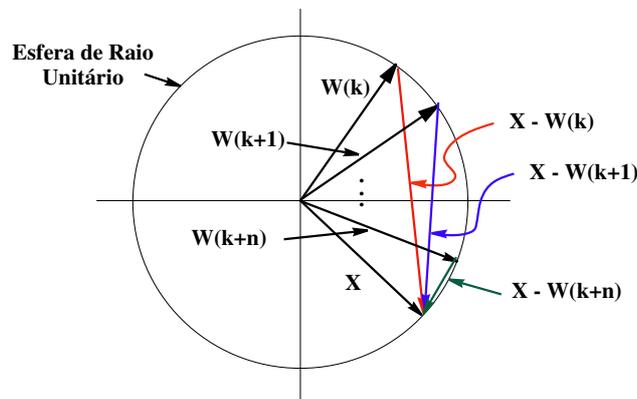


Figura 3.16: Aproximação de Vetores na Rede de Kohonen

Tanto a rede neural de kohonen quanto as demais técnicas apresentadas neste capítulo foram aplicadas ao problema proposto no Capítulo 2 e seus resultados são apresentados no próximo capítulo.

Capítulo 4

Implementação e Análise dos Sistemas de Diagnóstico Propostos

*E*ste capítulo é dedicado a apresentar os sistemas de diagnóstico desenvolvidos neste trabalho, bem como o procedimento de aplicação da norma ANSI/IEEE Std C57.104- (1991) ao conjunto de dados, fornecidos pela Companhia Energética de Minas Gerais (CEMIG), referente à análises cromatográficas realizadas em óleo isolante de seus transformadores de potência, para obtenção dos diagnósticos de falta nestes equipamentos. Estes dados foram utilizados nos processos de aprendizado dos sistemas desenvolvidos.

4.1 Base de Dados Utilizada

Conforme mencionado no Capítulo 1 a base de dados utilizada no desenvolvimento dos sistemas de diagnóstico propostos neste trabalho foi fornecida pela Companhia Energética de Minas Gerais (CEMIG) e era composta por informações relativas a 777 amostras de óleo isolante extraídas de transformadores de potência instalados em todo o estado de Minas Gerais, coletadas no período entre dezembro de 1993 até março de 2002. Estas informações constavam da concentração, em ppm, dos onze principais gases que normalmente se encontram dissolvidos em óleo isolante, conforme mostrado na Tabela 2.1, extraídos a partir de análises cromatográficas às quais estas amostras foram submetidas, sendo que o procedimento de análise cromatográfica foi executado segundo a norma ANSI/IEEE Std C57.104- (1978). Por serem considerados como irrelevantes no processo de diagnóstico (DÖRNENBURG e STRITTMATTER, 1974), (FALLOU, 1975), (ROGERS, 1975), o Propano (C_3H_8) e o Propileno (C_3H_6) foram desconsiderados desta base de dados.

Contudo, para a utilização dos dados desta base no desenvolvimento dos sistemas de diagnóstico necessitou-se não apenas das concentrações dos gases que se encontravam dissolvidos no óleo, mas também dos diagnósticos relativos a tais amostras, para que assim fosse possível se proceder o treinamento e ajuste dos sistemas desenvolvidos. Como não foi possível obter tais informações, foi desenvolvido um sistema que, a partir das informações das concentrações dos gases, pudesse simular um diagnóstico para cada uma das amostras da base de dados. Este sistema foi desenvolvido conforme normatização internacional, sendo apresentado na próxima seção.

4.2 Sistema de Simulação de Diagnóstico

Para que os procedimentos de treinamento dos sistemas de diagnóstico pudessem ser implementados, foi inicialmente necessário o desenvolvimento de um algoritmo que, a partir das informações referentes às análises cromatográficas, gerasse um diagnóstico, que seria tomado como referência para o sistema em desenvolvimento.

Desta forma, foi necessário adotar um dos critério de diagnóstico de faltas a partir da análise de gases dissolvidos, descritos no Capítulo 2, para que este servisse de base para o

sistema de simulação proposto.

Uma vez que o método dos gases chave e o método das razões de Dörnemburg diagnosticam apenas um limitado número de faltas (ver Capítulo 2) o método das razões de Rogers foi tomado como base deste sistema, sendo que este método é, também, o recomendado pela normalização internacional (ANSI/IEEE Std C57.104- 1991).

O sistema de simulação de diagnóstico com a finalidade de automatizar o método das razões de Rogers foi implementando segundo o fluxograma apresentado na Figura 4.1, que serviu de referência para a validação dos resultados obtidos pelos sistemas de diagnóstico propostos neste trabalho.

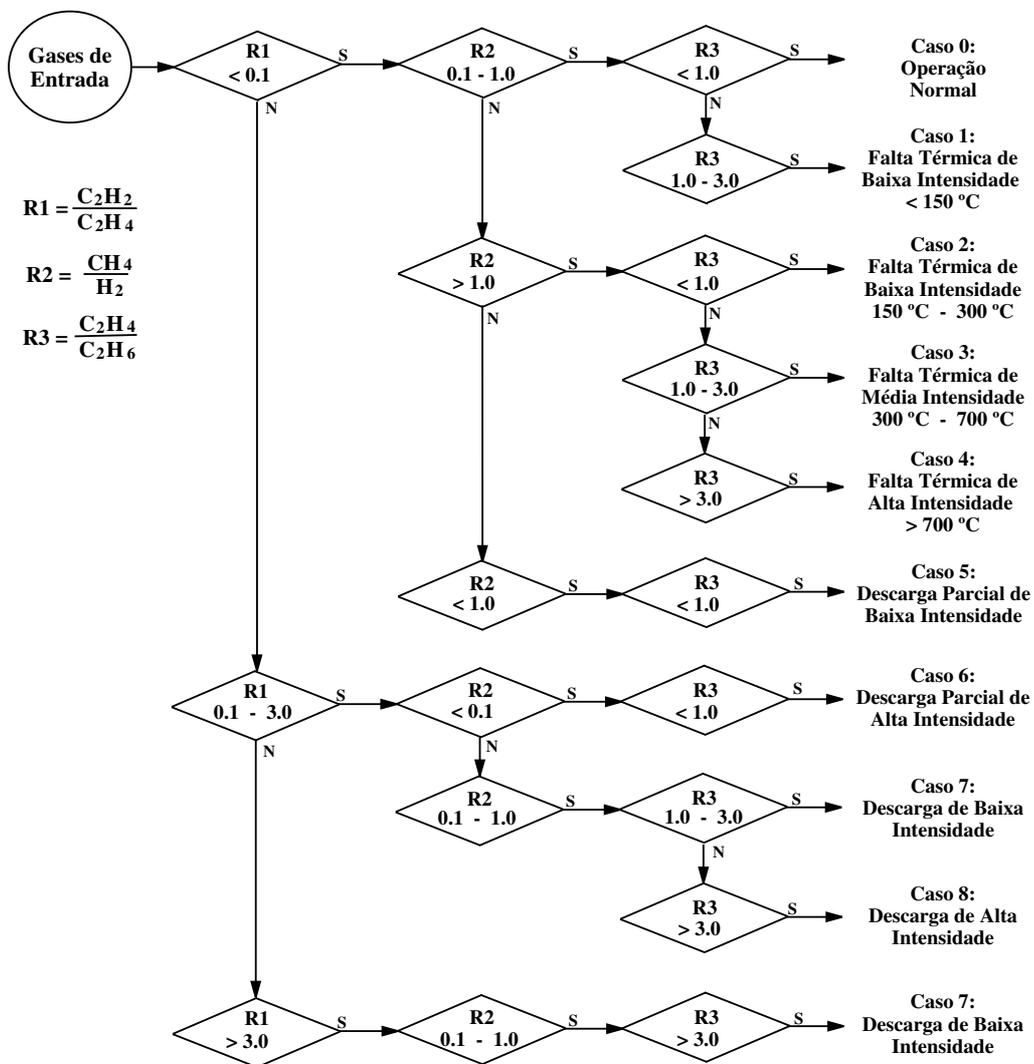


Figura 4.1: Fluxograma do Sistema de Simulação de Diagnóstico

4.3 Técnicas de Inteligência Computacional Aplicadas ao Diagnóstico de Falhas em Transformadores de Potência

Na tentativa de resolver o problema de identificação de falhas em transformadores de potência a partir da análise de gases dissolvidos em óleo isolante, foram utilizadas quatro técnicas de inteligência computacional, conforme apresentadas no Capítulo 3. As considerações de implementação de cada uma destas técnicas, assim como a comparação entre elas são discutidas a seguir.

4.4 Rede Neural MLP Aplicada no Diagnóstico de Falhas em Transformadores de Potência

A justificativa de obter-se uma estrutura de sistema de diagnóstico de falta em transformadores de potência usando uma rede neural MLP foi devido ao seu campo de aplicação bastante amplo e sua grande capacidade de generalização.

Conforme discutido na Seção 4.2, as entradas para o sistema de diagnóstico utilizando a rede MLP foram tomadas como sendo as relações de gases que fazem parte do método das Razões de Rogers, ou seja, $\frac{C_2H_2}{C_2H_4}$, $\frac{CH_4}{H_2}$ e $\frac{C_2H_4}{C_2H_6}$ e a saída aquelas obtidas através do sistema de simulação de diagnóstico da Figura 4.1.

O processo de treinamento, conforme mencionado no Capítulo 3, foi construído utilizando-se o algoritmo *backpropagation*. Este algoritmo de treinamento de Redes Neurais é um dos mais utilizados em aplicações práticas de previsão, classificação e reconhecimento de padrões em geral. Apesar do grande sucesso do *backpropagation*, foram verificados algumas dificuldades inerentes à sua aplicação durante o desenvolvimento deste trabalho. Estes problemas, bem como as soluções adotadas são apresentados a seguir:

Definição do tamanho da rede: Para se definir o tamanho de uma rede neural deve-se definir o número de camadas intermediárias e o número de neurônios em cada uma dessas camadas. Estes números devem conter um compromisso entre **convergência**, que é a capacidade da Rede Neural de aprender todos os padrões do conjunto de treinamento, e **generalização**, que é a capacidade de responder corretamente à padrões nunca vistos.

Se a rede neural for pequena, não será capaz de armazenar todos os padrões necessários, por outro lado se a rede for muito grande (muitos parâmetros = pesos), esta poderá não responder corretamente aos padrões nunca vistos.

Uma das maneiras de se resolver este problema é a utilização da métrica de Hecht-Nielsen (HECHT-NIELSEN, 1990), baseado no teorema de Kolmogorov-Nielsen (1957), que afirma que com apenas uma camada oculta já é possível calcular uma função arbitrária qualquer a partir de dados fornecidos. De acordo com esta métrica, a camada oculta deve ter por volta de $2n + 1$ neurônios, onde n é o número de variáveis de entrada.

Problema de paralisia da Rede Neural: Com o treinamento, os pesos (ω_{ji}) podem alcançar valores muito altos. Com isso, a soma ponderada de cada processador torna-se também muito alta, uma vez que $v_j^{(l)}(n) = \sum_{i=0}^m \omega_{ji}^{(\ell)}(n) y_i^{(\ell-1)}(n)$. Aplicando-se uma função de ativação logística não-linear $\varphi(v_j^{(l)}(n))$ à saída dos neurônios, obtém-se valores que aproximam-se da região de saturação da função de ativação (aproximadamente 1), conforme mostrado na Figura 4.2. Desta forma a derivada da função de ativação $\varphi'(v_j^{(l)}(n))$, e conseqüentemente a correção dos pesos (ver Equação (3.26)), tende para zero, fazendo com que a rede pare de "aprender".

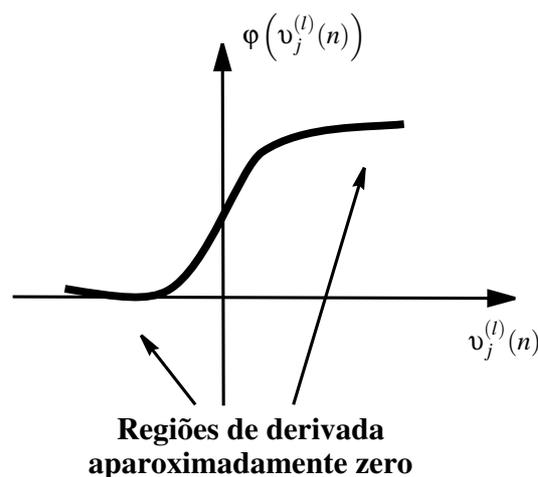


Figura 4.2: Exemplo de Função de Ativação

Para resolver o problema de paralisia da rede, é necessário escolher valores iniciais de

pesos e limiares uniformemente distribuídos dentro de um intervalo pequeno. Um intervalo muito utilizado na literatura para a inicialização dos pesos com a finalidade de evitar este problema é o intervalo entre $[-0.1, +0.1]$. Outro procedimento que visa evitar o problema de paralisia é o de não permitir que o valor das entradas da rede alcancem valores muito altos. Isso foi feito através de uma normalização, conforme descrito na Equação (4.1), fazendo com que todas as entradas estivessem contidas no intervalo $[0,1]$.

$$x_i = \frac{[R_{1_i}, R_{2_i}, R_{3_i}]^T}{\|[R_{1_i}, R_{2_i}, R_{3_i}]^T\|} \quad (4.1)$$

onde x_i é o i -ésimo padrão de entrada, $R_{1_i} = \frac{CH_{4_i}}{H_{2_i}}$, $R_{2_i} = \frac{C_2H_{2_i}}{C_2H_{4_i}}$ e $R_{3_i} = \frac{C_2H_{4_i}}{C_2H_{6_i}}$.

Problema de Mínimos Locais: Com a utilização de uma taxa de aprendizagem η muito pequena na Equação (3.26), não é possível calcular uma variação de pesos que faça a Rede Neural sair de um Mínimo Local, conforme ilustrado na Figura 4.3, não atingindo a performance desejada (erro mínimo a ser atingido).

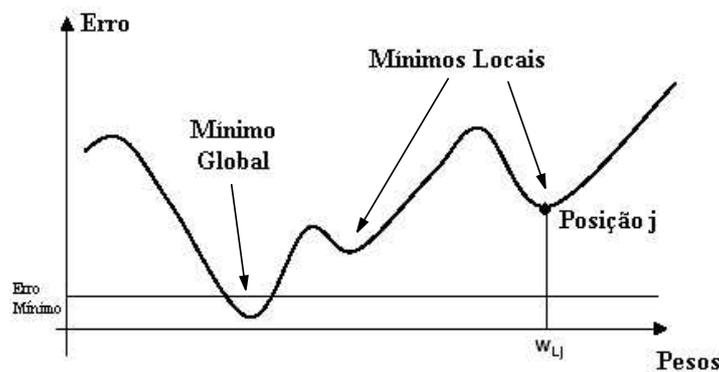


Figura 4.3: Problema de Mínimos Locais

Por outro lado, com a utilização de taxas de aprendizagem muito grandes, a Rede Neural pode nunca conseguir chegar ao Mínimo Global pois as variações nos pesos se tornam muito grandes, gerando instabilidade no processo de treinamento.

Para resolver este problema, é possível utilizar uma taxa de aprendizagem η adaptativa. Neste processo o valor de η é inicialmente fixado como um valor alto (próximo de 1) e depois reduzido gradativamente de acordo com a convergência do erro.

Desta forma, a estrutura da rede MLP a ser utilizada estava praticamente definida, faltando apenas o número de neurônio de saída. Este parâmetro foi definido empiricamente, através dos testes de desempenho de três possíveis estruturas.

Assim, a primeira destas estruturas a ser testada foi uma rede com três neurônios de entrada, 7 neurônios na camada Intermediária e apenas uma saída, conforme mostrado na Figura 4.4:

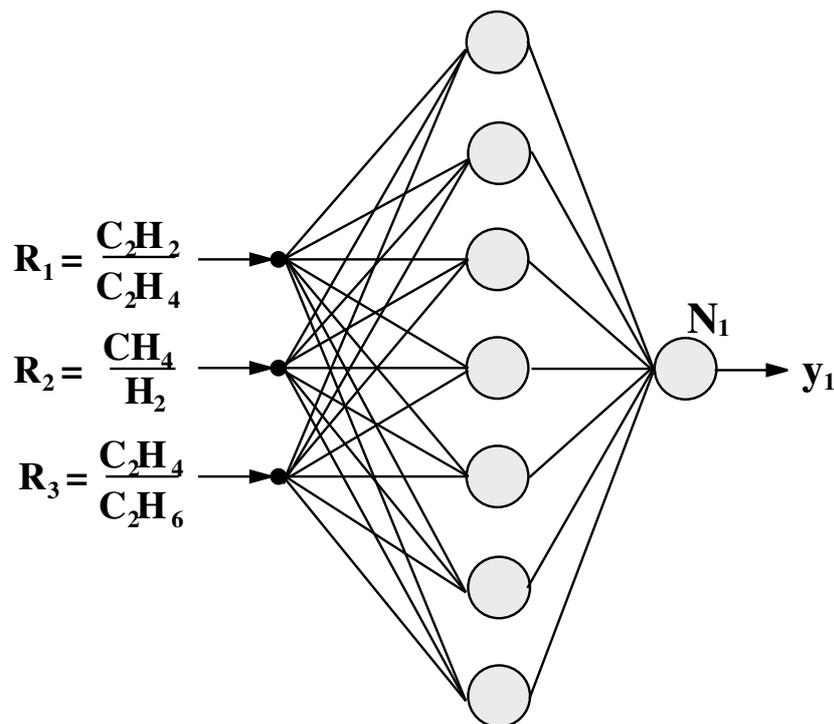


Figura 4.4: Primeira Estrutura de Rede Neural MLP Implementada

Nesta rede, todos os neurônios foram implementados usando uma função de transferência do tipo sigmoidal, apresentada na Equação 4.2 :

$$y_j^{(\ell)} = \frac{1}{1 + e^{-\delta \times v_j^{(\ell)}}} \quad (4.2)$$

onde $y_j^{(\ell)}$ é a saída do neurônio j da camada ℓ , $v_j^{(\ell)}$ é o mesmo definido na Equação 3.19, e δ representa a inclinação da função, conforme mostrado na Figura 4.5:

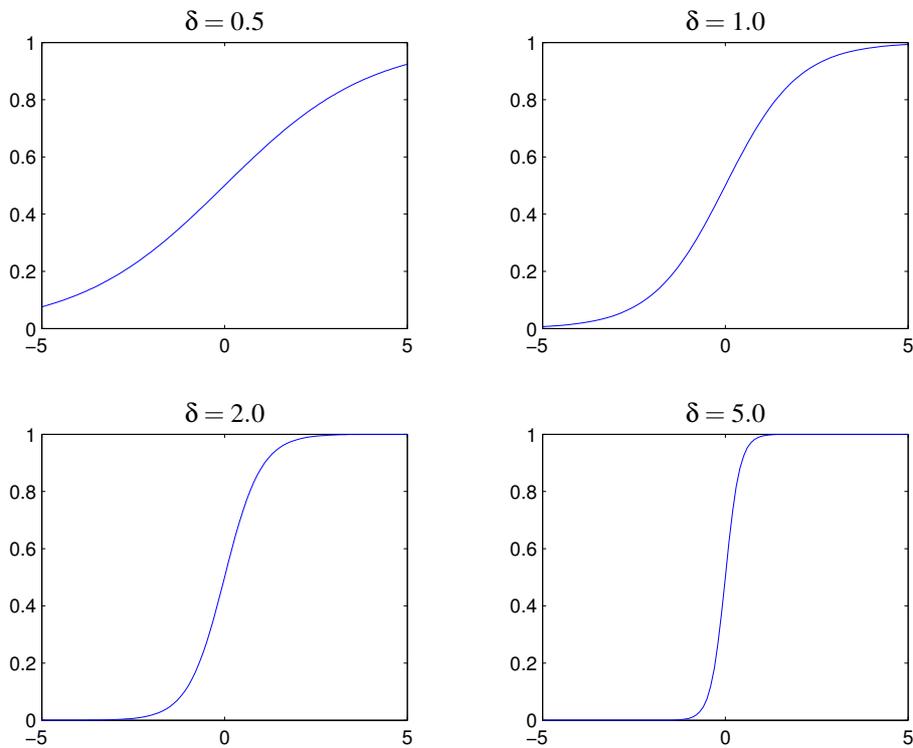


Figura 4.5: Função de Transferência Sigmoidal

Uma vez que o neurônio de saída y_1 também implementava uma função sigmoial, a saída da rede estava compreendida no intervalo entre $[0,1]$. Para relacionar a saída da rede com os códigos de faltas da Tabela 2.5 foi criada a Tabela 4.1.

Tabela 4.1: Correlação entre os Tipos de Falta e a Saída da Primeira Rede Neural MLP

Tipo de Falta	N_1
0	0.0
1	0.1
2	0.2
3	0.3
4	0.4
5	0.5
6	0.6
7	0.7
8	0.8

Desta forma a rede foi treinada com 80 amostras de treinamento (aproximadamente 10%), extraídas da base de dados de forma aleatória. Os parâmetros usados no treinamento foram os apresentados na Tabela 4.2.

Tabela 4.2: Parâmetros Usados na Primeira Rede Neural MLP

Parâmetro	Valor
Número de Camadas	3
Número de Neurônios na Camada 1	3
Número de Neurônios na Camada 2	7
Número de Neurônios na Camada 3	1
Tolerância de Erro	10^{-5}
Taxa de Aprendizado	Adaptativa
Coefficiente δ da Função de Ativação	1
Coefficiente de Momento	0.9

Esta rede, apesar de apresentar uma estrutura leve e com baixo custo de treinamento, não obteve grande sucesso na identificação dos padrões de falta, uma vez que seu reduzido número de conexões não permitiu armazenar todo o comportamento não linear do problema. Em seu treinamento, os valores de tolerância de erro não foram atingidos, apresentando assim altos índices de erro, quando aplicados à toda a base de dados, demonstrando também um baixo poder de generalização.

Com o fracasso da primeira rede, tentou-se uma segunda estrutura, com um maior número de conexões, a fim de aumentar a capacidade de armazenamento da rede. Desta forma foi proposta uma estrutura formada por 3 neurônios de entrada, 7 neurônios na camada intermediária e 4 neurônios de saída, conforme mostrado na Figura 4.6.

A camada de saída da rede neural MLP representa o código de diagnóstico do método das razões de Rogers apresentado na Tabela 2.5, porém de forma binária. Assim, para obter-se os vetores de treinamento utilizados no processo de aprendizado é necessário converter os códigos de falta relativos às amostras de entrada em números binários.

Da mesma forma, para se extrair o código de falta após a aplicação de um padrão de entrada na rede é necessário converter o número binário gerado na saída da rede em um

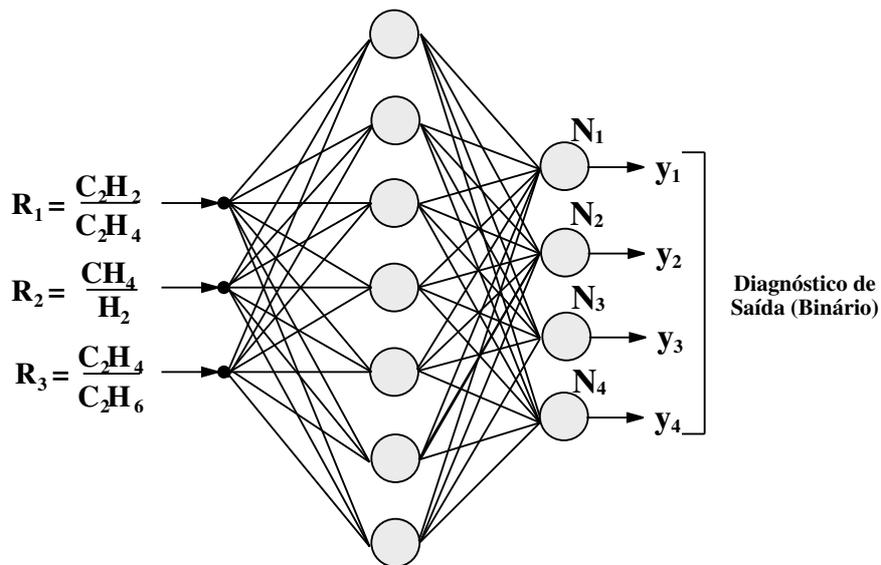


Figura 4.6: Estrutura da Segunda Rede Neural MLP Proposta

número decimal. Estas relações de correspondência entre a saída da rede neural MLP e o tipo de falta associado a ela pode ser visto na Tabela 4.3.

Tabela 4.3: Correlação entre os Tipos de Falta e a Saída da Rede Neural MLP

Tipo de Falta	N_1	N_2	N_3	N_4
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0

A configuração e os parâmetros usados na rede neural MLP e no algoritmo *backpropagation* são ilustrados na Tabela 4.4.

A função de ativação utilizada também foi a função sigmoideal, apresentada na Figura 4.5.

Para se obter o valor final na camada de saída, as saídas dos neurônios desta camada passam por um arredondamento, que toma como saída final o valor inteiro mais próximo da saída dos neurônios da camada de saída, garantindo que a saída da rede seja sempre binária.

Tabela 4.4: Parâmetros Usados na Rede Neural MLP

Parâmetro	Valor
Número de Camadas	3
Número de Neurônios na Camada 1	3
Número de Neurônios na Camada 2	7
Número de Neurônios na Camada 3	4
Tolerância de Erro	10^{-2}
Taxa de Aprendizado	Adaptativa
Coefficiente δ da Função de Ativação	1
Coefficiente de Momento	0.9

No processo de treinamento foram consideradas 70 amostras de treinamento, sendo que a entrada da rede é formada por um vetor 3×1 e a saída um vetor binário 4×1 . Os resultados dos testes realizados após o treinamento foram comparados aos obtidos segundo a norma ANSI/IEEE Std C57.104- (1991) e apresentados na Tabela 4.5.

Tabela 4.5: Resultados Obtidos para a Segunda Rede Neural MLP

Código de Falta	Número de Amostras Validadas	
	Diagnóstico IEEE	Diagnóstico MLP 2
0	34	24
1	42	39
2	209	205
3	25	23
4	122	117
5	1	1
6	20	20
7	142	137
8	182	177
TOTAL DE ACERTOS		743
PRECISÃO DE ACERTOS DA REDE		95,62%

Apesar da grande performance obtida por esta rede, tentou-se ainda uma terceira estrutura de rede, com a finalidade de melhorar ainda mais o desempenho do sistema de diagnóstico, bem como eliminar o decodificação das saídas da rede, o que diminuiria o custo computacional da rede.

Desta forma foi proposta uma estrutura composta por 3 neurônios de entrada, 7 neurônios na camada intermediária e 9 neurônios na camada de saída, onde cada um representava um

determinado tipo de falta. Esta estrutura é apresentada na Figura 4.7.

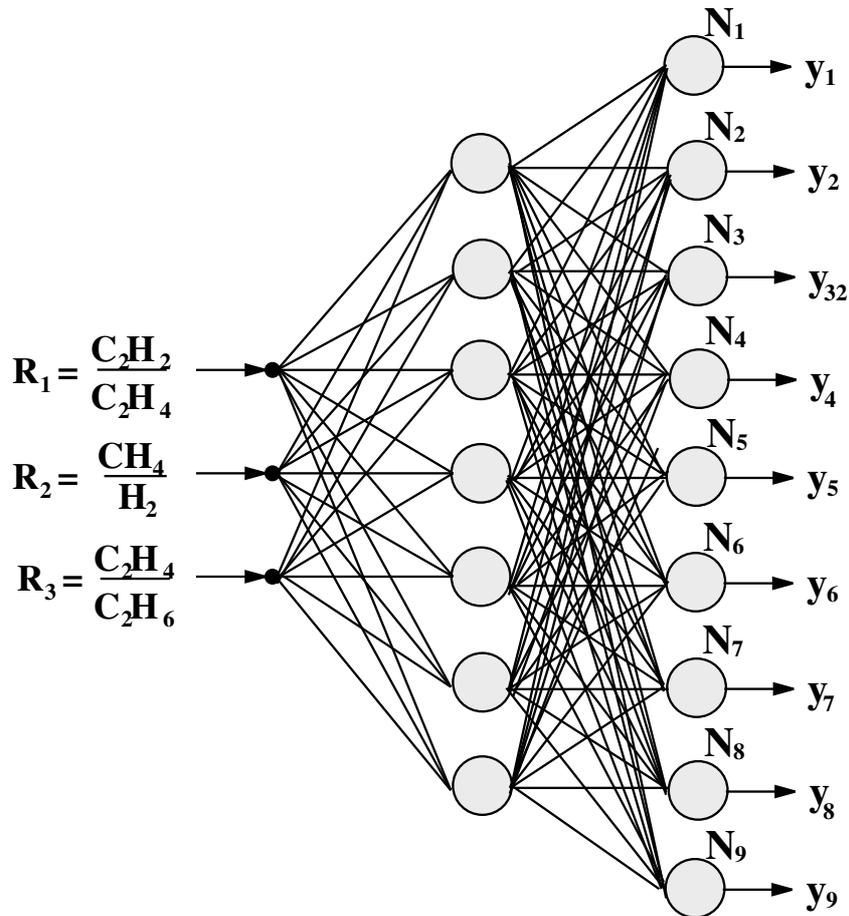


Figura 4.7: Estrutura da Terceira Rede Neural MLP Proposta

Da mesma forma que para as redes anteriores, nesta estrutura também foi utilizada a função de transferência sigmoideal e foram utilizadas 80 amostras escolhidas aleatoriamente em seu treinamento, sendo que os parâmetros da rede são os apresentados na Tabela 4.6.

A utilização desta rede apresentou vantagens tanto em relação a precisão de acertos quanto ao custo computacional, uma vez que não foi mais necessário decodificar uma saída binária para obter o diagnóstico. Neste caso a saída da rede passa apenas por uma função competitiva, e o neurônio que apresentar maior valor de saída é associado ao código de falta. Desta forma, quando a saída do neurônio N_1 vencia sobre as demais para uma determinada amostra, o código de diagnóstico relativa a esta amostra era igual a 0. Se o neurônio vencedor fosse o N_2 , o diagnóstico para a dada amostra seria 1, e assim por diante.

Tabela 4.6: Parâmetros Usados na Rede Neural MLP

Parâmetro	Valor
Número de Camadas	3
Número de Neurônios na Camada 1	3
Número de Neurônios na Camada 2	7
Número de Neurônios na Camada 3	9
Tolerância de Erro	10^{-2}
Taxa de Aprendizado	Adaptativa
Coefficiente δ da Função de Ativação	1
Coefficiente de Momento	0.9

Os resultados obtidos por esta estrutura são apresentados na Tabela 4.7.

Tabela 4.7: Resultados Obtidos para a Segunda Rede Neural MLP

Código de Falta	Número de Amostras Validadas	
	Diagnóstico IEEE	Diagnóstico MLP 3
0	34	34
1	42	42
2	209	209
3	25	25
4	122	122
5	1	0
6	20	0
7	142	137
8	182	182
TOTAL DE ACERTOS		751
PRECISÃO DE ACERTOS DA REDE		96,65%

Outras estruturas também foram testadas, tais como redes com um menor número de neurônios na camada intermediária e com maior número de entradas (levando-se em consideração a concentração absoluta dos gases, e não as relações). Entretanto o desempenho destas redes não superou, tanto em precisão como em custo, os das redes aqui apresentadas.

Levando-se em consideração seu desempenho, a terceira rede foi considerada, dentre as redes MLP testadas, como estrutura que melhor resolveu o problema em questão e foi selecionada para a comparação com as demais sistemas apresentados a seguir.

4.5 Rede Neural de Função de Base Radial Aplicada no Diagnóstico de Falhas em Transformadores de Potência

O uso deste tipo de técnica no diagnóstico de falhas em transformadores de potência é motivado pelo seu rápido e efetivo processo de treinamento, que garante o reconhecimento de algumas características do conjunto de treinamento mais facilmente, se comparado com uma rede perceptron de múltiplas camadas treinada com *backpropagation*.

Da mesma forma que para a rede neural MLP, as entradas para a rede neural RBFN são as três relações de gases relativas ao método das razões de Rogers.

A definição dos campos receptivos para esta rede foi obtida através de procedimento heurístico. A partir de um número inicial de campos receptivos a complexidade da rede foi sendo aumentada e os parâmetros e a precisão da rede foram sendo avaliadas a cada passo. A estrutura ótima da rede foi definida com base no compromisso entre a complexidade e a precisão do sistema. Esta estrutura é apresentada na Figura 4.8.

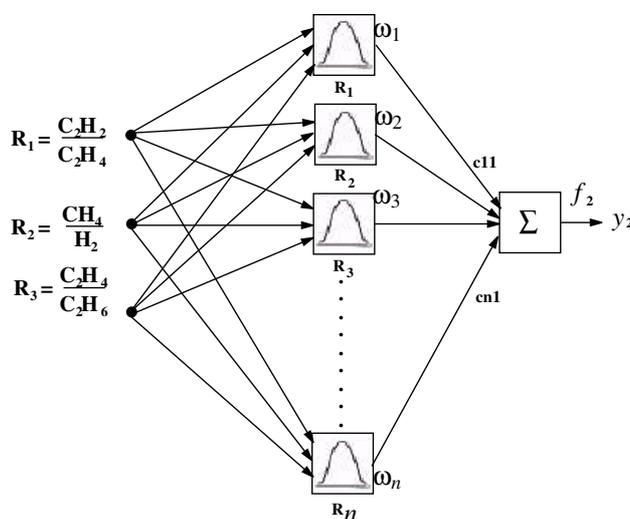


Figura 4.8: Estrutura da Rede Neural RBFN Proposta

A identificação dos parâmetros da rede é feita a cada passo através de um 'aprendizado híbrido' (JANG, 1997), no qual os parâmetros dos campos receptivos (u_i , σ_i e c_i) são inicialmente fixados pelo método de auto organização LVQ (*Learning Vector Quantization*) (KOHONEN, 1991), somente para as regiões do espaço de entrada onde existam dados im-

portantes e, então, os pesos da camada de saída são ajustados através do algoritmo ortogonal de mínimos quadrados (CHEN et. al., 1991).

A RBFN proposta foi treinada com as mesmas amostras utilizadas na treinamento da rede neural MLP. Após o treinamento foram identificados 72 centros c_i ($n = 72$). Entretanto o desempenho da rede quando treinado com as mesmas 80 amostras anteriores foi bem inferior do que o apresentado pela rede MLP. Contudo, quando aumentamos o número de amostras de treinamento para 160 (aproximadamente 20% da base), escolhidas aleatoriamente, o número de centros encontrados sobe para 157, fazendo com que os resultados tivessem uma significativa melhoria. Os Resultados finais para a rede neural RBFN são apresentados na Tabela 4.8.

Tabela 4.8: Resultados Obtidos para a Rede Neural RBFN

Código de Falta	Número de Amostras Validadas	
	Diagnóstico IEEE	Diagnóstico RBFN
0	34	34
1	42	42
2	209	207
3	25	25
4	122	120
5	1	1
6	20	20
7	142	138
8	182	174
TOTAL DE ACERTOS		761
PRECISÃO DE ACERTOS DA REDE		97,94%

4.6 Modelo de Kohonen Aplicado no Diagnóstico de Falhas em Transformadores de Potência

A utilização da rede neural de Kohonen para obter-se uma estrutura de diagnóstico teve como justificativa sua alta capacidade de classificação de padrões e baixo custo computacional, uma vez que o processo de diagnóstico de falhas com base em um vetor de relações de gases dissolvidos pode ser visto como um problema clássico de classificação de padrões, onde cada classe refere-se a um tipo de falta.

Inicialmente, como entrada para a rede neural de Kohonen, foram utilizadas as mesmas variáveis de entrada utilizadas nos sistemas de diagnóstico apresentados anteriormente, ou

seja, as três razões de Rogers, resultando na estrutura apresentada na Figura 4.9.

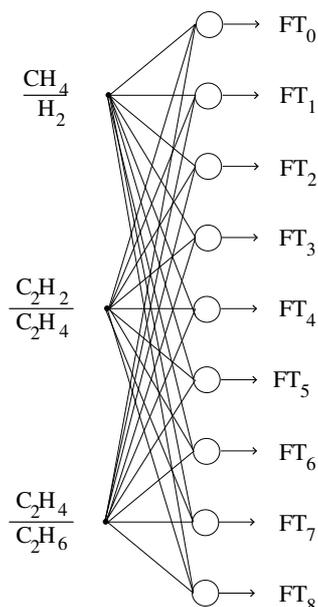


Figura 4.9: Estrutura da Camada de Kohonen Inicialmente Proposta

Esta estrutura foi composta por três nós de entrada, referentes às três relações de gases do método das razões de Rogers, e nove neurônios de saída, sendo que cada um deles está relacionado com um tipo de falta listados na Tabela 2.5.

Entretanto, quando aplicada ao problema proposto, a rede não conseguiu classificar corretamente os padrões de entrada em seus respectivos grupos de falta de forma adequada. Verificou-se que, devido ao fato de a rede de Kohonen identificar padrões pertencentes a mesmos grupos através da distância Euclidiana entre estes padrões, isto tornava sua aplicação ineficaz no caso específico, pois padrões de entrada (transformados em vetores no espaço Euclidiano) mesmo muito distantes no espaço poderiam pertencer ao mesmo grupo de falta, enquanto que padrões relativamente próximos poderiam pertencer a grupos diferentes. Isso se dá pois pelo método das Razões de Rogers, uma vez excedidos determinados limites das razões, por maior que fossem os valores destas razões o diagnóstico resultante não se alterava, fazendo com que vetores muito distantes tenham o mesmo diagnóstico, enquanto que vetores que se encontrem próximos aos limites, mesmo com uma pequena distância entre eles, pudessem pertencer a grupos diferentes.

A forma encontrada de solucionar este problema foi fuzificar os padrões de entrada tomando-se como base os conjuntos fuzzy apresentados nas Figuras 2.2, 2.3 e 2.4. Desta forma, os vetores que excedessem os limites mencionados acima teriam pertinência igual a 1 no conjunto alto, independente do quanto esta razão excedeu seu valor limite. Assim, a nova estrutura testada foi composta por nove neurônios na camada de entrada, sendo que cada um deles representava a pertinência, para cada uma das 3 razões de entrada, nos conjuntos Baixo ($\mu_L(\cdot)$), Médio ($\mu_M(\cdot)$) e Alto ($\mu_H(\cdot)$). A saída, da mesma forma que na estrutura anterior, era composta por nove neurônios, cada um representando um tipo de falta. Esta estrutura é ilustrada na Figura 4.10.

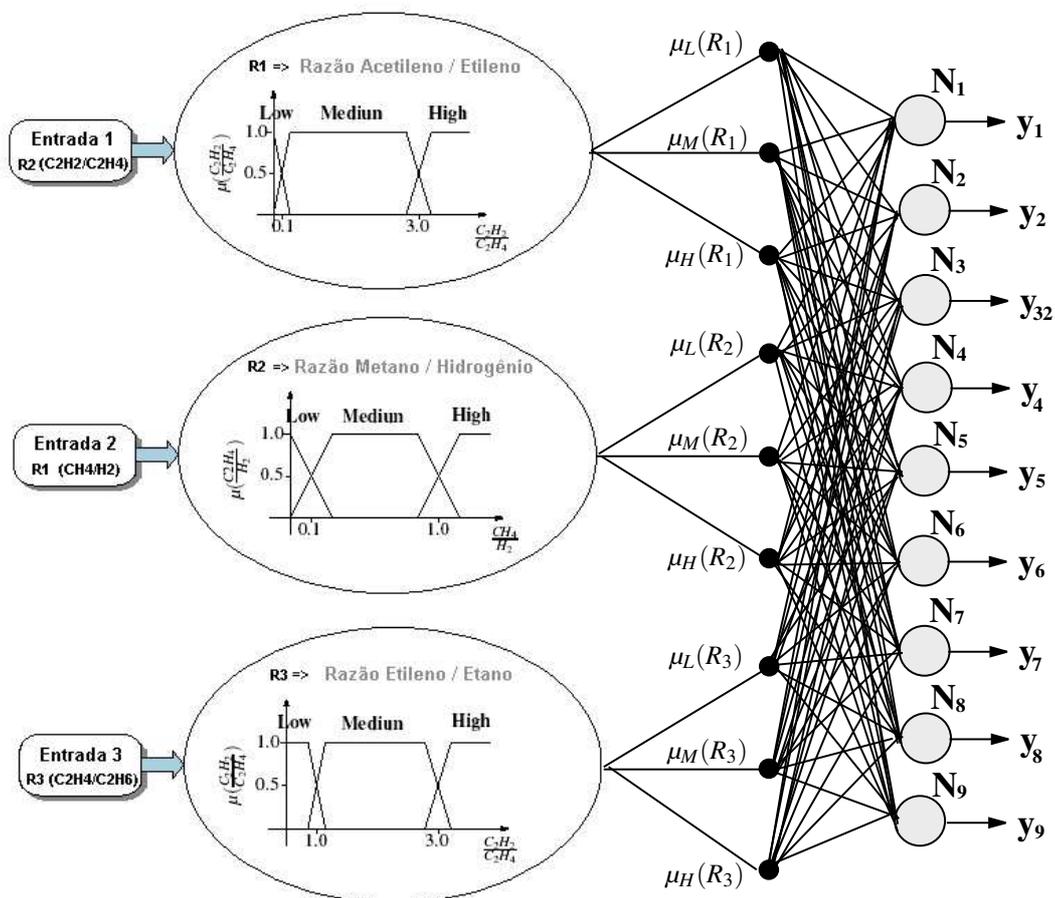


Figura 4.10: Estrutura da Rede Neural de Kohonen Proposta

Como este tipo de rede neural não tem capacidade de generalização, ou seja, nesta rede, apresentando-se um padrão, este irá ativar um único neurônio de saída, representando um

tipo de falta, toda a base de dados é utilizada no processo de aprendizado não-supervisionado implementado para este tipo de sistema.

Desta forma, o processo de identificação do tipo de falta referente às amostras de análises cromatográficas em óleo isolante é iniciado com o cálculo da menor distâncias entre os padrões de entrada fuzificados $X = [\mu_L(R_1), \mu_M(R_1), \mu_H(R_1), \mu_L(R_2), \mu_M(R_2), \mu_H(R_2), \mu_L(R_3), \mu_M(R_3), \mu_H(R_3)]$ e os vetores de peso associados a cada um dos 9 neurônios que compõem a saída da rede neural. Assim, o neurônio relativo ao vetor de pesos mais próximo do vetor de entrada X é considerado como vencedor e o tipo de falta referente à amostra X será aquele associado ao neurônio vencedor.

Como resultado é possível apresentar a alocação dos pontos e seus respectivos conjuntos em um espaço tridimensional, levando-se em consideração os valores das razões de Rogers, conforme apresentado na Figura 4.11. Neste gráfico, cada um dos pontos plotados corresponde a uma amostra de relações de gases, e cada um dos formatos possíveis de serem assumidos por estes pontos correspondem a um tipo de falta, conforme mostrado na legenda.

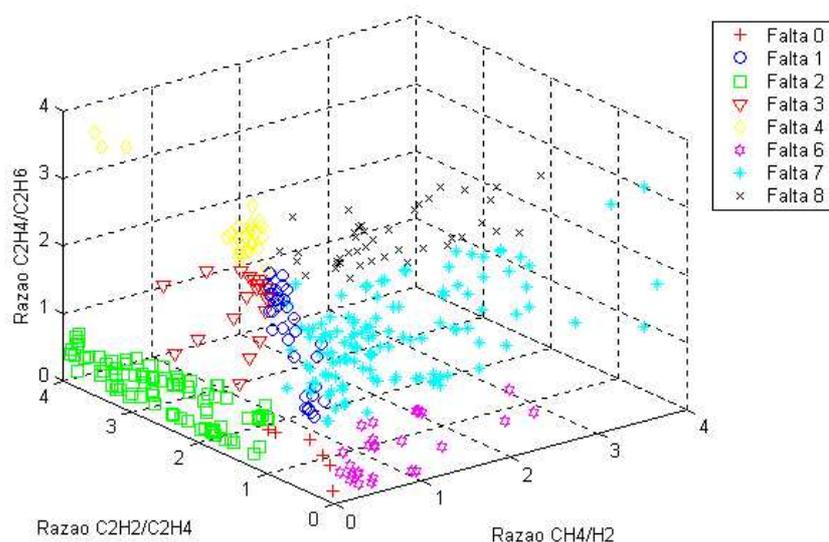


Figura 4.11: Resultados Obtidos para o Sistema de Diagnóstico Baseado na Rede Neural de Kohonen

Além disso, os resultados numéricos referentes à aplicação da rede neural de Kohonen são apresentados na Tabela 4.9.

Tabela 4.9: Resultados Numéricos Obtidos para o Sistema de Diagnóstico Baseado na Rede Neural de Kohonen

Código de Falta	Número de Amostas Validadas	
	Diagnóstico IEEE	Diagnóstico Kohonen
0	34	34
1	42	41
2	209	208
3	25	25
4	122	121
5	1	1
6	20	20
7	142	140
8	182	180
TOTAL DE ACERTOS		770
PRECISÃO DE ACERTOS DA REDE		99,10%

4.6.1 ANFIS Aplicado no Diagnóstico de Falhas em Transformadores de Potência

A justificativa da implementação desta técnica no problema proposto foi devido à significativa melhoria na eficiência do sistema de inferência fuzzy descrito por (DUKARM, 1993) para um conjunto específico de equipamentos.

Da mesma forma que nos sistemas anteriores, para o ANFIS as entradas foram tomadas como sendo as três razões de Rogers apresentadas anteriormente.

Assim, inicialmente foi construído um sistema de inferência fuzzy (FIS) de Sugeno de ordem zero, conforme descrito em (DUKARM, 1993). Este sistema é apresentado na Figura 4.12.

Os dados de entrada, compreendidos por 777 amostras de análises cromatográficas em óleo isolante, foram então aplicados ao sistema de FIS desenvolvido, obtendo-se os resultados apresentados na Tabela 4.10.

Após esse passo, uma estrutura ANFIS, conforme descrito na Seção 3.3, foi desenvolvida para ajustar a forma dos conjuntos fuzzy e aumentar a precisão global do sistema para o conjunto de dados considerado, conforme apresentado em (HELL et al., 2002c). Esta estrutura representa um sistema de inferência fuzzy de Sugeno de ordem zero. Como neste caso a saída da rede não é uma função, conforme apresentado na Seção 3.3, e sim conjuntos *singleton* fuzzy, os parâmetros consequentes não são ajustados e a saída é obtida aplicando-se uma função competitiva aos neurônios da Camada 4, conforme apresentado na Figura 4.13.

No processo de treinamento do sistema foram utilizadas 80 amostras de treinamento

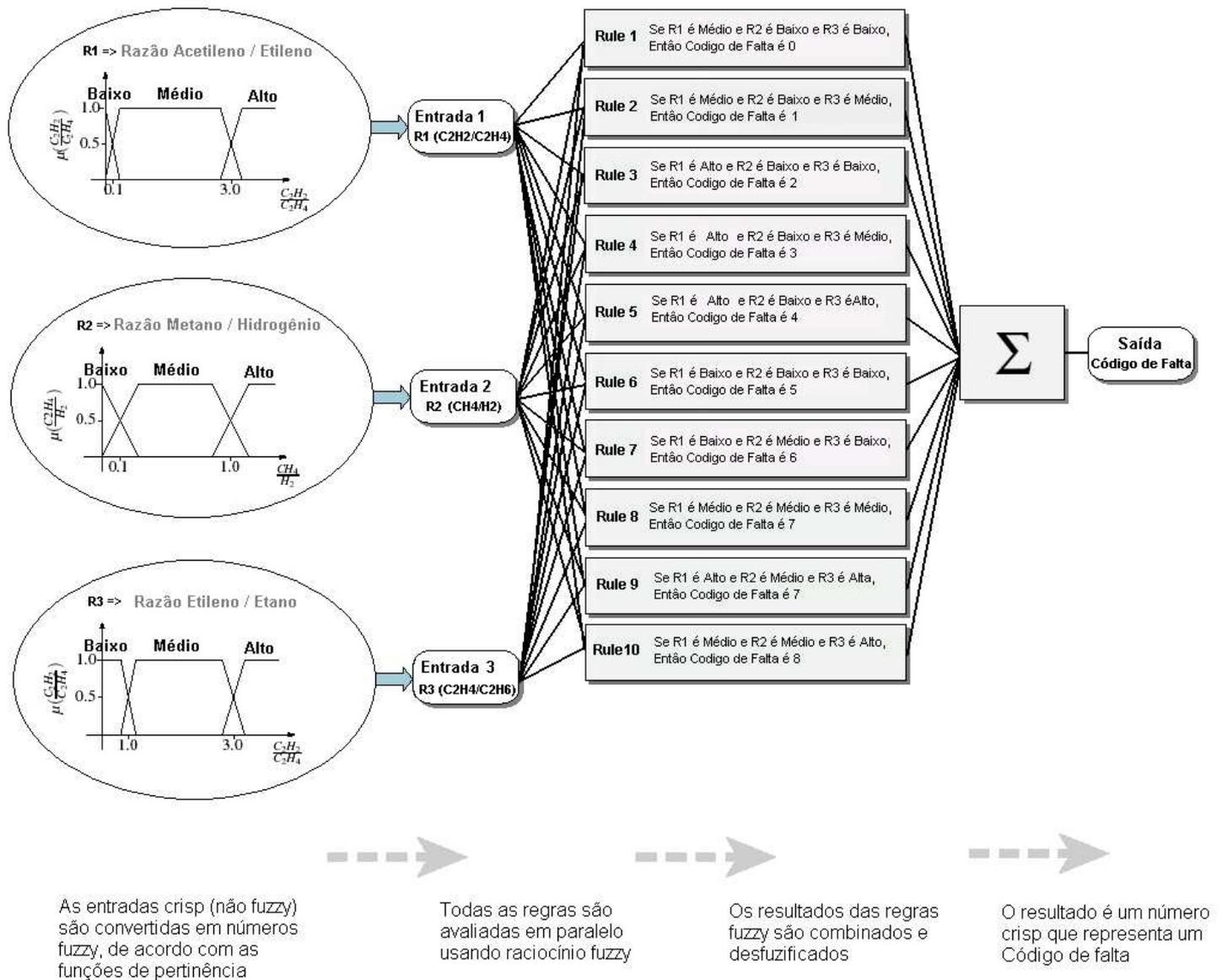


Figura 4.12: Sistema de Inferência Fuzzy Implementado (FIS)

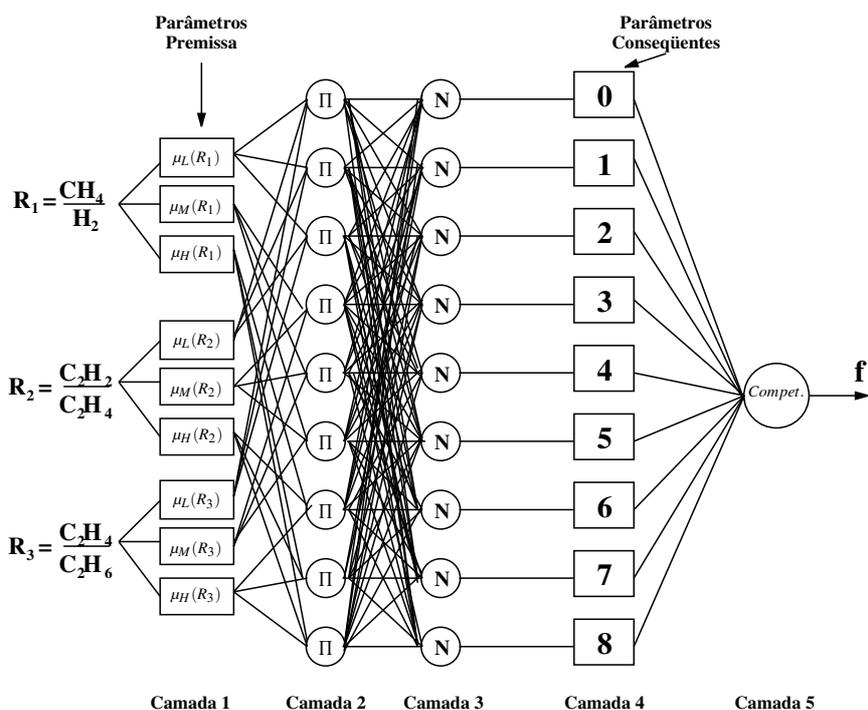


Figura 4.13: Estrutura ANFIS Proposta

(aproximadamente 10% da base de dados) e os testes de validação foram feitos tomando-se todas as amostras da bases, inclusive as usadas no treinamento. Os resultados obtidos para o sistema ANFIS também são apresentados na Tabela 4.10.

Tabela 4.10: Resultados Obtidos para o Sistema de Diagnóstico ANFIS

Código de Falta	Número de Amostras Validadas		
	Diagnóstico IEEE	Diagnóstico FIS	Diagnóstico ANFIS
0	34	34	34
1	42	40	42
2	209	208	209
3	25	23	25
4	122	121	122
5	1	1	1
6	20	20	20
7	142	127	142
8	182	168	182
TOTAL DE ACERTOS		742	777
PRECISÃO DE ACERTOS DA REDE		95,50%	100,00%

4.7 Análise Comparativa dos Resultados Apresentados pelos Sistemas de Diagnóstico de Falhas em Transformadores de Potência

Apesar de um menor número de amostras de treinamento terem sido usadas no processo de aprendizado das redes neurais MLP e nas redes neurais de função de base radial se comparados ao modelo de Kohonen, estas apresentaram erros significativos durante a fase de teste e validação. Porém, uma vez que os diagnósticos usados no treinamento foram obtidos a partir da simulação baseada na norma internacional (ANSI/IEEE Std C57.104- 1991), muitas vezes não totalmente aplicados à realidade da planta de equipamentos analisada devido às condições climáticas aos quais estes equipamentos estão submetidos, é possível que estas técnicas tenham melhor aproximação para diagnósticos reais obtidos em campo, devido a sua grande capacidade de mapeamento não linear, sem ainda necessitar de conhecimentos *a priori* do problema, como é o caso do sistema ANFIS.

O sistema baseado no modelo de Kohonen obteve grande sucesso na fase de testes, obtendo erros muito pequenos se comparados com os padrões obtidos pela simulação da norma. O uso deste sistema requer que todos os vetores que compõem a base de dados analisada sejam utilizados nos procedimentos de treinamento e teste, porém este problema é minimizado pelo fato do algoritmo de treinamento ser extremamente rápido, superando todos os demais sistemas em termos de custos computacionais.

O sistema ANFIS alcançou o melhor ajuste dentre todos os sistemas propostos, reconhecendo todos os padrões de falta de forma correta. No processo de treinamento deste sistema foi exigido um menor custo computacional se comparado às redes neurais MLP e às redes neurais de função de base radial, uma vez que foram necessários um menor número de amostras de treinamento no ajuste do espaço de saída.

A partir desta análise, nota-se que o sistema baseado no modelo de Kohonen, mesmo apresentado pequenos índices de erro no diagnóstico, ainda é o mais indicado para aplicações com esta característica, uma vez estes índices de erro são totalmente aceitáveis.

A Tabela 4.11 apresenta um resumo dos resultados obtidos para os sistemas de diagnóstico propostos.

Tabela 4.11: Resumo dos Resultados Numéricos Obtidos

Código de Falta	Número de Amostras Validadas						
	Diagnóstico IEEE	Diagnóstico MLP 2	Diagnóstico MLP 3	Diagnóstico RBFN	Diagnóstico Kohonen	Diagnóstico FIS	Diagnóstico ANFIS
0	34	24	34	34	34	34	34
1	42	39	42	42	41	40	42
2	209	205	209	207	208	208	209
3	25	23	25	25	25	23	25
4	122	117	122	120	121	121	122
5	1	1	0	1	1	1	1
6	20	20	0	20	20	20	20
7	142	137	137	138	140	127	142
8	182	177	182	174	180	168	182
TOTAL DE ACERTOS		743	751	761	770	742	777
PRECISÃO DE ACERTOS DA REDE		95,62%	96,65%	97,94%	99,10%	95,50%	100,00%

Vale ressaltar mais uma vez que durante para obtenção dos resultados da Tabela 4.11 os dados relativos aos diagnósticos foram obtidos através da aplicação da norma ANSI/IEEE Std C57.104- (1991). Contudo, pretende-se ainda obter os reais tipos de faltas relativos às amostras testadas, com a finalidade de validar a aplicação dos sistemas desenvolvidos para fins reais de diagnóstico em campo. Durante as primeiras tentativas de obtenção destes diagnósticos, verificou-se que certos equipamentos, mesmo contendo altas concentrações de gases dissolvidos no óleo isolante, não apresentavam nenhum tipo de falta, mesmo que incipiente, enquanto que outros, contendo baixas concentrações de gases no óleo, não estavam livres de faltas.

Isso se dá devido à grande diversidade de tipos de equipamentos da qual é composta a planta analisada, formada por transformadores de vários níveis de tensão e potência, tipos de fabricante, projeto, óleo isolante, tempos de operação. Estes fatores influenciam no processo de geração de gases, causando muitas vezes alarmes falsos, ou não permitindo que uma deterioração do material isolante seja detectada, comprometendo a operação do equipamento.

Para resolver este problema é proposto no Anexo 5.2 um sistema de pré-processamento da base de dados utilizada. Esta técnica permite que os sistemas propostos, mesmo tendo sido desenvolvidos a partir de dados de diagnóstico simulados, tenham desempenho similar quando aplicados ao problema real de diagnóstico.

Capítulo 5

Conclusões e Propostas de Trabalhos Futuros

Neste capítulo são apresentadas as análises dos resultados mais importantes do trabalho, bem como a proposição de trabalhos futuros.

5.1 Conclusões

Este trabalho concentrou-se na solução do problema de diagnóstico de faltas em transformadores de potência com base na análise dos gases que se encontram dissolvidos no óleo isolante.

Inicialmente, demonstrou-se o processo de classificação da base de dados usada no desenvolvimento dos sistemas de diagnóstico propostos. Este processo se mostrou compatível com a realidade do conjunto de equipamentos tomados como exemplo neste trabalho.

Após este primeiro passo, mostrou-se como os diagnósticos utilizados como saída desejada nos processos de treinamento foram obtidos, tomando-se como base a normalização internacional (ANSI/IEEE Std C57.104- 1991).

Apresentou-se, também, quatro técnicas de inteligência computacional, e discutiu-se suas capacidades de reconhecimento de faltas, chegando a conclusão que os sistemas de inferência fuzzy neuro-adaptativos e os modelos de Kohonen possuem grande precisão na solução do problema proposto, sendo que o modelo de Kohonen requer menores custos computacionais, se comparado ao sistema ANFIS.

Vale mais uma vez salientar que os diagnósticos utilizados nos processos de treinamento e validação dos sistemas desenvolvidos foram obtidos a partir da aplicação de norma internacional, a que muitas vezes não se adequa à realidade brasileira. Um objetivo futuro deste trabalho é empregar os sistemas de diagnóstico propostos tomando-se como saídas desejadas diagnósticos reais de falta, colhidos em campo. Neste caso específico é possível que outra técnica se mostre mais adequado do que o modelo de Kohonen obteve para os diagnósticos gerados pela Norma (ANSI/IEEE Std C57.104- 1991).

Exemplos numéricos foram apresentados, verificando a eficiência dos sistemas de diagnóstico propostos para o diagnóstico de faltas em transformadores de potência com base na análise de gases dissolvidos.

5.2 Proposição para Trabalhos Futuros

Todos os sistemas de diagnóstico propostos neste trabalho foram desenvolvidos para trabalhar em modo *off-line*, ou seja, para que um diagnóstico possa ser obtido por qualquer um destes

sistemas, é necessário que um operador humano entre com os dados referentes às análises cromatográficas em óleo isolante via dispositivo de entrada externo. Uma linha de continuidade deste trabalho seria a implementação destes sistemas de diagnóstico em modo *on-line*, uma vez que é de interesse de empresas do setor elétrico nacional e internacional, a instalação de sensores remotos para obtenção dos dados referentes aos gases que se encontram dissolvidos no óleo. Desta forma seria possível monitorar de forma constante os transformadores de potência, aumentando ainda mais a confiabilidade do sistema.

Outra linha de continuidade deste trabalho seria a implementação de um processo de gerenciamento de diversos sistemas especialistas para diagnósticos de faltas em equipamentos elétricos, de forma que fosse possível agrupar em um único sistema as ferramentas de diagnóstico desenvolvidas neste trabalho com outras, desenvolvidas para resolver problemas de diagnóstico em outros tipos de equipamentos, tais como em máquinas síncronas, assíncronas e turbo-geradores, fornecendo um poderoso instrumento de suporte à manutenção de equipamentos elétricos de um modo geral.

Uma terceira linha de continuidade seria o desenvolvimento de um modelo térmico de transformadores de potência que tome como base informações como corrente de carga, temperatura de topo de óleo, condições climáticas, etc., com a finalidade de obter as temperaturas internas de operação destes transformadores, conforme ilustrado na Figura 5.1.



Figura 5.1: Esquema Básico de Modelo Térmico de Transformadores de Potência

Com base neste modelo e no desenvolvimento de mais dois outros modelos, um que tome como base as temperaturas internas dos transformadores e estime a vida útil residual do equipamento, e outro que simule o comportamento termodinâmico do óleo isolante, computando os

gases que estão sendo gerados a partir das temperaturas internas do equipamento (ilustrados na Figura 5.2), é possível obter tanto estimativa mais precisa das condições de sobrecarga máxima permissível aos transformadores, quanto informações a respeito dos gases gerados no interior dos mesmos.



Figura 5.2: Esquema Básico de Modelo Térmico de Transformadores de Potência

Desta forma é possível relacionar estes modelos com os sistemas de diagnóstico propostos neste trabalho para a obtenção de dados relativos a possíveis faltas nos transformadores, sem a necessidade das caras análises cromatográficas, conforme ilustrado na Figura 5.3.



Figura 5.3: Esquema Básico de Modelo Térmico de Transformadores de Potência

Bibliografia

- ANSI/IEEE Std C57.104-1978. IEEE Guide for the detection and determination of generated gases in oil-immersed transformers and their relation to the serviceability of the equipment **American National Standard - ANSI**, 1978. 31p.
- ANSI/IEEE Std C57.104-1991. IEEE Guide for Interpretation of Gases Generated in oil-immersed Transformers **IEEE Power Engineering Society**, 2002. 45p.
- BEZDEK, C. J.; PAL, S. K. **Fuzzy Models for Pattern Recognition**. NY - New York-USA: IEEE Press, 1992. 539p.
- BRANCO FILHO, G. **Dicionário de termos de manutenção, confiabilidade e qualidade**. Rio de Janeiro-Brasil: abramam: Associação Brasileira de Manutenção , 1996. 122p.
- BACKER, E.; KANDEL, A. **Computer Assisted Reasoning in Cluster Analysis**. NY - New York-USA: Prentice Hall, 1995. 400p.
- CHEN, S.; COWAN, C. F. N.; GRANT, P. M. Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks. **IEEE Transaction on Neural Networks**, v.2, n.2, p.302-09, março 1991.
- DÖRNENBURG, E.; STRITTMATTER, W. Monitoring oil-cooled transformers by gas analysis. **Brown Boveri Review**, v.5, n.61, p.238-247, maio 1974.
- DUKARM, J. J. Transformer oil diagnosis using fuzzy logic and neural networks. In: **Canadian Conference on Electrical and Computer Engineering**, v.1, p.329-38, 1993.
- DUVAL, M. Dissolved Gas Analysis: It can save your transformer. **IEEE Electrical Insulation Magazine**, v.7, n.6, p.22-27, nov./dec. 1989.
- FALLOU, B.; VIALE, F.; DAVIES, I.; RORGERS, R. R.; DÖRNENBURG, E. Application of physico-chemical methods of analysis to the study of deterioration in the insulation of electrical apparatus. **CIGRE Report 15-07**, 1970.
- FALLOU, B. Detection of and research for the characteristics of an incipient fault from analysis of dissolved gases in the oil of an insulation. **Electra**, v.42, p.31-42, out. 1975.
- FOGEL, D. B. An introduction to simulated evolutionary optimization. **IEEE Transaction on Neural Networks**, v.1, n.5, p.3-14, jan. 1994.
- GALDI, V.; IPPOLITO, L.; PICCOLO, A.; VACCARO, A. Diagnostic System for Transformer Thermal Overload Protection. **Electric Power Applications, IEE Proceedings**, v.147, n.5, p.415-421, set. 2000.

- HALSTEAD, W. D. A thermodynamic assessment of the formation of gaseous hydrocarbons in faulty transformers. **Journal of the Institute of Petroleum**, v.59, n.569, p.239-41, set. 1973.
- HAYKIN, S. **Neural Networks : A Comprehensive Foundation**. Englewood Cliffs - NJ-USA: Macmillan College Publishing Company, Inc., 1994. 696p.
- HAYKIN, S. **Redes Neurais : Princípios e Práticas**. 2 ed. Porto Alegre-RS - Brasil: Bookman, 2001. 900p.
- HECHT-NIELSEN, R. **Neurocomputing**. NY - New York-USA: Addison Wesley Publisher, 1990. 433p.
- HELL, M. B.; D'ANGELO, M. F. S. V.; COSTA Jr., P. P. Power Transformer Fault Diagnosis Based on Dissolved Gas Analysis Using a Fuzzy Neural Network Approach in a Real Data Base. In: ZEMLIAK, A.; MASTORAKIS, E. **Advances in Information Science and Soft Computing**, NY-USA: WEAS Press, 2002. Cap.34, p.181-186.
- HELL, M. B.; D'ANGELO, M. F. S. V.; SILVA, N. A.; FONSECA, G. I. B. da; COSTA Jr., P. P. Power Transformer Fault Diagnosis Based on Dissolved Gas Analysis Using a Kohonen Neural Network Approach in a Real Data Base. In: **V INDUSCON - IEEE International Conference on Industry Applications**, 2002, Salvador-BA.
- HELL, M. B.; D'ANGELO, M. F. S. V.; SILVA, N. A.; FONSECA, G. I. B. da; COSTA Jr., P. P. An Adaptative Neuro-Fuzzy Inference System Based on Dissolved Gas Analysis for Fault Diagnosis in Power Transformers. In: **2002 WSEAS International Conference on Power Engineering Systems**, 2002, Rio de Janeiro-RJ.
- HOWE, V. H.; MASSEY, L.; WILSON, A. M. Identity and significance of gases collected in Buchholz protectors. **IM. V. Gazette**, maio 1956.
- HUANG, Y. C.; YANG, H. T.; HUANG, C. L. Developing a new transformer fault diagnosis system through evolutionary fuzzy logic. **IEEE Transaction on Power Delivery**, v.12, n.2, p.761-67, abril 1997.
- JANG, J.-S. R. ANFIS: Adaptative-Network-Based Fuzzy Inference System. **IEEE Transactions on System, Man, and Cybernetics**, v.23, n.3, p.665-685, maio/junho 1993.
- JANG, J.-S. R.; SUM, C.-T.; MIZUTANI, E. **Neuro-Fuzzy and Soft Computing**. NY-USA: Prentice Hall, 1997. 614p. (Matlab Curriculum Series)
- KELLY, J. J. Transformer fault diagnosis by dissolved-gas analysis. **IEEE Transaction on Electric Insulation**, v.16, n.6, p.777-82, nov./dec. 1980.
- KOHONEN, T. The Self Organizing Map. In: LAU, C. **Neural Networks: Theoretical Foundations and Analysis**, NY-USA: IEEE Press, 1991. p.181-186.
- LUTTRELL, S. P. Self-supervised training of hierarchical vector quantizers. In: **2nd International Conference on Artificial Neural Networks**, 1991, Bournemouth, England. p.5-9.
- MANDAMI, E. H.; ASSILIAN, S. An experiment in linguistic synthesis with a fuzzy logic controller. **International Journal of Man-Machine Studies**, v.7, n.1, p.1-13, 1973.

- RAMIREZ-RODRIGUEZ, C.; VLADIMIROVA, T. Fuzzy Neural Network for Fuzzy Classification. **IEEE International Conference on Systems, Man and Cybernetics, Intelligent Systems for the 21st Century**, v.1, p.322-327, fev. 1995.
- ROGERS, R. R. UK Experience in the interpretation of incipient faults in power transformers by dissolved gas-in-oil chromatographic analysis. **Doble Client Conference, Paper 42 AIC 75**, 1975.
- ROGERS, R. R. IEEE and IEC codes to interpret incipient faults in transformers using gas-in-oil analysis. **IEEE Transaction on Electric Insulation**, v.El-13, n.5, p.348-54, out. 1978.
- SIMPSON, P. K. Fuzzy Min-Max Neural Networks - Part 2: Clustering. **IEEE Transaction on Fuzzy Systems**, v.1, n.1, p.32-45, fev. 1993.
- SPUDAT, R. W. Manutenção Preventiva em Transformadores. **Revista Manutenção**, v.1, p.17-21, maio/junho 1993.
- TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling control. **IEEE Transactions on System, Man, and Cybernetics**, v.15, p.116-132, 1985.
- TSOUKALAS, L. H.; UHRIG, R. E. **Fuzzy and Neural Approaches in Engineering (Adaptive and Learning Systems for Signal Processing, Communications and Control Series)**. NY-USA: John Wiley and Sons, Inc., 1997. 587p.
- TSUKAMOTO, Y. An approach to fuzzy reasoning method. In: GUPTA, M. M.; RAGADE, R. K.; YAGER, R. R. **Advances in fuzzy set theory and applications**, Amsterdam: North-Holland, 1979. p.137-149.
- WANG, Z.; LIU, Y.; GRIFFIN, P. J. Neural net and expert system diagnose transformer faults. **IEEE Computer Applications in Power**, v.13, n.1, p.50-55, Jan. 2000.
- YANG, H. T.; HUANG, Y. C. Intelligent decision support for diagnosis of incipiente transformer faults using self-organizing polinomial networks. **IEEE Transaction on Power Systems**, v.13, n.3, p.946-52, agosto 1998.
- ZADEH, L. A. Fuzzy logic. **IEEE Computer**, v.21, n.4, p.83-93, abril 1988.
- ZANG, Y.; DING, X.; LIU, Y.; GRIFFIN, P. J. An artificial neural network approach to transformer fault diagnosis. **IEEE Transaction on Power Delivery**, v.11, n.4, p.1836-41, jan. 1996.

Abstract

In this work the problem of diagnosis incipient faults in power transformers through the dissolved gas analysis in the insulation oil was treated. In the solution of the problem was used five techniques of computational intelligence to develop diagnosis systems that could recognize the faults presented in the analyzed transformers. The first developed system was based on a multi layer Perceptron neural network, trained with backpropagation algorithm. The second developed system was based on a radial basis function neural network approach. The third applied technique resulted in the development of a Kohonen neural network. The fourth developed system was based on a fuzzy inference system and the fifth system implemented was a adaptive neuro fuzzy inference system. From the obtained results was possible conclude that the Kohonen neural network is the most indicated system to classify faults in power transformers. The work also presented an application of Adeli-Hung algorithm to cluster the analyzed transformers, with the purpose of classify the utilized data base to obtain better results when these results will be validate by real diagnosis collected in field.

Método de Pré-processamento da Base de Dados

5.2.1 Agrupamento da Base de Dados Utilizada

Com a finalidade de tratar o problema apresentado no Capítulo 4, neste anexo é proposto um sistema de pré-processamento do conjunto de dados para que os equipamentos fossem agrupados em classes (ou famílias) de equipamentos e pudessem ser, então, tratados de forma diferenciada.

Devido à sua grande simplicidade de implementação e de seus bons resultados, o algoritmo escolhido para o agrupamento dos dados de entrada foi o algoritmo de Adeli-Hung (TSOUKALAS e UHRIG, 1997), conforme apresentado a seguir:

O Algoritmo de Adeli-Hung

O algoritmo de de Adeli-Hung (TSOUKALAS e UHRIG, 1997) é desenvolvido usando uma topologia de rede de duas camadas onde o número de nós de entrada é igual aos M atributos de cada um dos padrões de entrada a serem agrupados e o número de nós de saída igual ao número de conjuntos aos quais os dados de entrada devem ser associados.

Inicialmente, este algoritmo usa uma rede neural com M entradas e uma saída. Quando o primeiro padrão é apresentado à rede, este fica associado ao primeiro *cluster* (representado pelo primeiro neurônio da rede). Se o segundo padrão apresentado é classificado para o primeiro *cluster*, o nó de saída representando o primeiro cluster torna-se ativo. Se o segundo exemplo de treinamento é classificado como um novo *cluster*, um nó de saída é adicionado à rede, e assim por diante, até que todos os exemplos sejam treinados.

Desenvolvimento do Algoritmo

Segundo (TSOUKALAS e UHRIG, 1997), para desenvolver o algoritmo de Adeli-Hung, uma função $diff(X, C_i)$, denominada *grau de diferença* é definida. Esta função representa a diferença entre exemplo de treinamento e um *cluster*, definido por seu vetor centro C_j , para $j = 1, \dots, P$, onde P indica o número de *clusters*. Esta função mapeia dois vetores dados (X e C_j) à um número real. Os vetores centro de cada *clusters* (média dos padrões dos exemplos nos *cluster*) são armazenados nos pesos da rede.

Procedimento

O procedimento apresentado a seguir permite classificar um exemplo de treinamento num cluster ativo ou novo utilizando o algoritmo de Adeli-Hung:

Passo 1: Calcule o grau de diferença, $diff(X, C_j)$, para $j = 1, \dots, P$, entre o exemplo de treinamento X , e cada um dos P *clusters*. A distância Euclidiana é usada para obter a função $diff(X, C_j)$, isto é:

$$diff(X, C_j) = \sqrt{\sum_{i=1}^M (X_i - C_{ji})^2} \quad (5.1)$$

Passo 2: Calcular o menor grau de diferença, $\min(diff(X, C_j))$, e fazer o cluster com menor grau de diferença um cluster ativo:

$$C_{ativo} = \{C^* / C^* = \arg \min(diff(X, C_j)), j = 1, 2, \dots, P\} \quad (5.2)$$

Passo3 Compare o valor de $\min(diff(X, C_j))$ com um valor limite predefinido d . Se o valor de $\min(diff(X, C_j))$ é maior que d , o exemplo é classificado como um novo cluster (Neste ponto mais um nó de saída é criado).

$$C_{novo} = X \quad \text{se} \quad \exists j | d < \min(diff(X, C_j)), j = 1, 2, \dots, P \quad (5.3)$$

Para entender melhor, suponha que N exemplos de treinamento tenham sido organizados em P grupos. Se os clusters são disjuntos, cada exemplo no conjunto de treinamento pertence à somente um dos clusters e uma matriz Z pode ser usada para representar o cluster associado a cada exemplo. Se o exemplo i pertence ao cluster j , então $Z_{ij} = 1$, caso contrário $Z_{ij} = 0$. Se por outro lado os clusters se sobrepõe, um dado exemplo de treinamento pode ser compatível com mais de um cluster. Portanto as fronteiras dos clusters classificados tornam-se fuzzy em vez de crisp.

Uma mesma matriz Z pode ser usada para representar o cluster de cada exemplo. O centro para cada cluster é definido como a média de todos os elementos do cluster, enquanto que o grau de pertinência de cada exemplo no cluster é definido a partir de sua similaridade com o centro.

A similaridade é definida como uma função da distância entre o exemplo e o centro do cluster. Se existem N exemplos num cluster P onde o k -ésimo elemento é $X_k^P = [x_{k1}^P, x_{k2}^P, \dots, x_{kM}^P]$, então o centro C_P (a média de todos exemplos) do cluster P é definido por:

$$C_P = [c_{p1}, c_{p2}, \dots, c_{pM}] = \frac{1}{N} \sum_{k=1}^N X_k^P \quad (5.4)$$

ou seja, $c_{pi} = \frac{1}{N} \sum_{k=1}^N x_{ki}^P$ e $i = 1, \dots, M$.

Usando uma função de pertinência triangular, o valor de pertinência do k -ésimo exemplo no j -ésimo cluster é definido como:

$$\mu_j(X_k^j) = f \left[D^W (X_k^j, C_j) \right] = \begin{cases} 0 & \text{se } D^W (X_k^j, C_j) > \delta \\ 1 - \frac{D^W (X_k^j, C_j)}{\delta} & \text{se } D^W (X_k^j, C_j) \leq \delta \end{cases} \quad (5.5)$$

onde um valor predefinido de limiar δ é usado como valor de crossover. A função similaridade é definida como uma norma ponderada $D^W (X_k^j, C_j)$. A norma ponderada no algoritmo de Adeli-Hung é a distância Euclidiana, da forma:

$$D^W (X_k^j, C_j) = \|w_j (X_k^j, C_j)\|^W = \sqrt{\sum_{i=1}^M (x_{ki}^j - c_{ji})^2} \quad (5.6)$$

onde W e w_j usualmente recebem valor 1. Se a distância Euclidiana é menor do que o valor do limiar δ , o exemplo pertence ao cluster j para um dado grau de pertinência.

5.3 Agrupamento da Base de Dados Utilizando o Algoritmo de Adeli-Hung

Conforme mencionado anteriormente, preparar os sistemas de diagnóstico para se adaptarem a diagnósticos reais obtidos em campo o algoritmo de Adeli-Hung foi implementado um sistema para pré-processar o conjunto de dados, agrupando os equipamentos em classes (ou famílias) de equipamentos.

Assim, a cada transformador foi associado um vetor de atributos, contendo todas as características referentes a estes equipamentos. Conforme citado na Seção 5.2.1, inicialmente apenas uma classe foi definida juntamente com um raio de vizinhança d , onde o primeiro vetor atributo foi tomado como sendo o centro desta classe. A medida que os vetores atributo X foram sendo apresentados ao sistema, a distância entre estes vetores e o vetor centro C_i das classes existentes ($diff(X, C_i)$) eram calculadas e comparadas com o raio d , sendo que o equipamento relativo a este vetor é atribuído à classe com menor valor de $diff(X, C_i)$, desde que este valor não ultrapassasse d . Caso isso ocorresse, uma nova família de transformadores era criada, tomando, neste instante, o vetor X como sendo o centro desta nova classe.

Desta forma foram definidas 7 famílias de transformadores, sendo que este número de classes se aproximou bastante do número de famílias de transformadores consideradas pela Companhia Energética de Minas Gerais, classificadas a partir da experiência de diversos anos de operação destes equipamentos.

Uma vez agrupados, os equipamentos analisados podem ser tratados de formas diferenciadas, dependendo da família a que pertencem, sendo que será possível o ajuste de um sistema de diagnóstico, seja ele qualquer um dos quatro tipos apresentados anteriormente, para cada um dos grupos obtidos. Esta idéia é apresentada na Figura 5.4.

Nesta estrutura, um determinado padrão de entrada será avaliado apenas pelo sistema de diagnóstico treinado para a família a qual pertence o equipamento que gerou este padrão, sendo que a saída dos demais sistemas de diagnóstico será igual a zero.

Com isso, pretende-se obter resultados ainda mais satisfatórios, quando os sistemas de diagnóstico forem testados com dados de diagnósticos reais extraídos em campo.

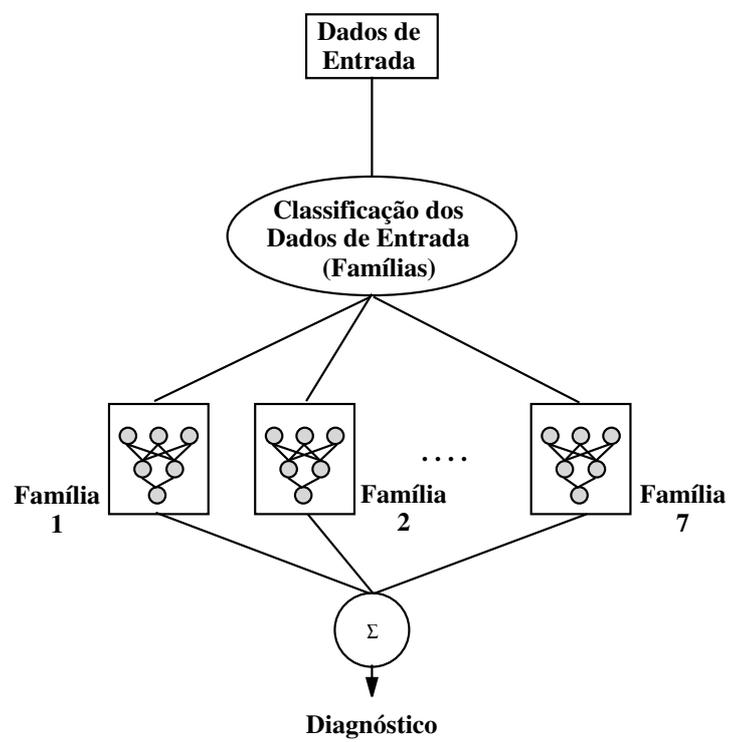


Figura 5.4: Avaliação por Famílias