



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
FACULDADE DE ENGENHARIAS ELÉTRICA E BIOMÉDICA

DANIEL DANTAS DO AMARAL RAMOS

**SISTEMA INTELIGENTE PARA DIAGNÓSTICO DE FALTAS EM
TRANSFORMADORES BASEADO EM ESTRUTURA DE REDES
NEURAIS AUTOASSOCIATIVAS EMPILHADAS**

BELÉM
2021

DANIEL DANTAS DO AMARAL RAMOS

**SISTEMA INTELIGENTE PARA DIAGNÓSTICO DE FALTAS EM
TRANSFORMADORES BASEADO EM ESTRUTURA DE REDES
NEURAIS AUTOASSOCIATIVAS EMPILHADAS**

Trabalho de Conclusão de Curso apresentado a Faculdade de Engenharias Elétrica e Biomédica do Instituto de Tecnologia da Universidade Federal do Pará, como requisito para a obtenção do Grau de Bacharel em Engenharia Elétrica.

Orientadora: Profa. Dra. Adriana Rosa Garcez Castro

BELÉM
2021

DANIEL DANTAS DO AMARAL RAMOS

**SISTEMA INTELIGENTE PARA DIAGNÓSTICO DE FALTAS EM
TRANSFORMADORES BASEADO EM ESTRUTURA DE REDES
NEURAIS AUTOASSOCIATIVAS EMPILHADAS**

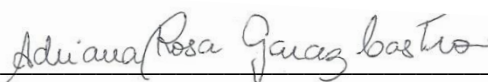
Trabalho de Conclusão de Curso apresentado a Faculdade de Engenharias Elétrica e Biomédica do Instituto de Tecnologia da Universidade Federal do Pará, como requisito para a obtenção do Grau de Bacharel em Engenharia Elétrica.

Orientadora: Profa. Dra. Adriana Rosa Garcez Castro

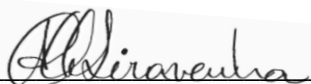
Data da avaliação: 7/10/2021

Conceito: EXCELENTE

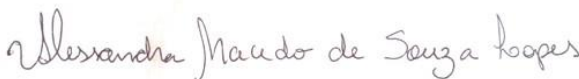
BANCA EXAMINADORA:



Profa. Dra. Adriana Rosa Garcez Castro
(FEEB/ITEC/UFPA – Orientadora)

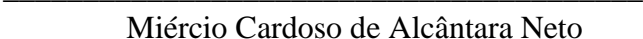


Profa. Dra. Ana Carolina Quintão Siravenha
(FEEB/ITEC/UFPA – Membro da Banca)



Profa. Dra. Alessandra Macedo de Souza Lopes
(ITEC/UFPA – Membro da Banca)

VISTO:



Miércio Cardoso de Alcântara Neto

Dedico esta monografia aos meus maiores
incentivadores, Ismael, Maria do Socorro,
Lívia, Diógenes e Angélica.

AGRADECIMENTOS

Agradeço aos meus pais, Ismael e Maria do Socorro, por me proporcionarem sempre as melhores condições e amor incondicional na construção da pessoa que sou hoje, à minha irmã Lívia que me suporta e apoia na vida cotidiana, pela amizade e companheirismo e à minha segunda mãe Ana, por todo carinho despendido a mim, vocês são fundamentais na minha vida. Adicionalmente, agradeço aos meus tios e padrinhos Diógenes e Dione por todo incentivo e cuidado durante os anos.

Agradeço aos meus amigos de faculdade e de vida por todo incentivo e camaradagem durante os anos, certamente a vida com vocês é mais alegre: Pedro, Amanda, Thiago, Mateus, Isac, Pavão, Rafael, Ádrea, Lauro, Gabriel, Jorge, João, Arthur, Iran e Felipe. Agradeço em especial à minha namorada Angélica Alcântara, a quem tive a sorte, honra e privilégio de conhecer na UFPA e a quem devo muitas das minhas felicidades e risos cotidianos desde então, obrigado por todo companheirismo, amor, amizade e por me ensinar todo dia como a vida pode ser bela e leve.

Agradeço aos mestres que me apoiaram e ensinaram muito do que sei hoje no âmbito profissional e pessoal. Professor Roberto Menezes, de quem tive a sorte de ser aluno em sua primeira turma na docência, obrigado pelo incentivo, conversas e conhecimentos repassados sobre o profissional e pessoal. Professor Claudomiro Barbosa, obrigado pela camaradagem, ensinamentos e oportunidades dadas a mim na graduação. Professores da escola Mendara, onde cresci e formei meu caráter, obrigado por tudo. Agradeço especialmente à minha orientadora Adriana Castro por todo apoio, paciência e orientações na reta final de minha trajetória como graduando.

Finalmente, agradeço a todas as boas pessoas a quem tive a sorte de conhecer e conviver na minha vida pessoal e acadêmica, todos meus familiares, tios e primos queridos: Marília, Calil, Lia, Maria Clara, Bianca, André, Lorena, Leo, Letícia e Rodrigo, e àqueles a quem o espaço não me possibilita citar todos os nomes, mas que com certeza contribuíram para este momento.

Portanto, a todos os que me inspiraram e ajudaram de alguma forma, meus sinceros agradecimentos.

“O pessimista reclama do vento, o otimista espera que ele mude, o realista ajusta as velas”

(Desconhecido)

RESUMO

Transformadores são equipamentos importantes e valiosos de qualquer sistema elétrico por proverem uma energia estável e confiável. A retirada de um transformador para manutenção, devido à falha interna, pode acarretar grande impacto para o sistema de energia sendo, portanto, fundamental um programa de manutenção preditiva. Métodos baseados na análise dos gases dissolvidos em óleo vêm sendo empregados pelas concessionárias de energia para o diagnóstico de falhas incipientes nos transformadores, todavia estes métodos sofrem ainda do problema de não-decisão, onde não se consegue obter um diagnóstico para determinadas combinações de concentração de gases, sendo necessário então a utilização de vários métodos para um diagnóstico final. Nesse cenário, este trabalho apresenta uma proposta de um sistema inteligente baseado em Redes Neurais Autoassociativas empilhadas para o diagnóstico de faltas em transformadores. O sistema proposto é baseado na análise cromatográfica dos gases dissolvidos em óleo isolante e utiliza a base de dados de transformadores faltosos TC 10, adicionalmente a outros dados provenientes da literatura, para o treinamento e teste do sistema. O resultado alcançado é promissor (acurácia de 100% na base de teste), considerando o sistema com apenas duas razões entre os gases dissolvidos em óleo como entrada. O resultado obtido mostra a aplicabilidade das Redes Neurais Autoassociativas ao problema em questão, portando-se então como uma boa alternativa para complementação, ou até substituição, aos métodos tradicionais utilizados pelas concessionárias.

Palavras-chave: Rede Neural Autoassociativa; análise cromatográfica; Diagnóstico de faltas; Inteligência computacional.

ABSTRACT

Transformers are important and valuable equipments in any electrical system for providing stable and reliable power. The removal of a transformer for maintenance, due to internal failures, can have a great impact on the power system, therefore, a predictive maintenance program is essential. Methods based on the analysis of gases dissolved in oil (DGA) have been used by power distribution companies to diagnose incipient transformer failures. However, these methods still suffer from the problem of non-decision, where it is not possible to obtain a diagnosis for certain combinations of gas concentrations, therefore being necessary more than one method to reach a final result. In this scenario, this work presents a proposal for an intelligent system based on stacked Autoassociative Neural Networks for transformer fault diagnosis. The proposed system is based on chromatographic analysis of gases dissolved in insulating oil and uses the TC 10 database of faulty transformers, in addition to other data from the literature, for training and testing the system. The result achieved is promising (accuracy of 100% on a test basis) considering the system with only two ratios of gases dissolved in oil as input. The result obtained shows the applicability of autoassociative neural networks to the problem in question, thus being a good alternative to complement, or even replace, the traditional methods used by companies.

Keywords: Autoassociative Neural Network; chromatographic analysis; Fault diagnosis; Computational Intelligence.

SUMÁRIO

1 INTRODUÇÃO	11
1.1 Objetivo geral	12
1.2 Objetivos específicos	12
1.3 Organização do trabalho	13
2 FUNDAMENTAÇÃO TEÓRICA	14
2.1 Diagnóstico de faltas em transformadores	14
2.1.1 Método de Dornenberg.....	18
2.1.2 Método de Rogers	19
2.1.3 Método IEC 60599	20
2.1.4 Método Duval.....	22
2.1.5 Método do Gás Chave	23
2.2 Redes Neurais Artificiais	23
2.2.1 Neurônio Artificial	24
2.2.2 Funções de ativação	25
2.2.2.1 Função Degrau	26
2.2.2.2 Função Linear.....	26
2.2.2.3 Função Sigmoidal.....	26
2.2.2.4 Função Tangente Hiperbólica	26
2.2.3 Arquiteturas de RNA.....	27
2.2.3.1 Redes alimentadas adiante de uma única camada	27
2.2.3.2 Redes alimentadas adiante com Múltiplas Camadas.....	28
2.2.3.3 Redes com realimentação.....	29
2.2.4 Paradigmas de Aprendizagem	30
2.2.4.1 Algoritmo Backpropagation	31
2.2.5 Generalização em RNA.....	32
2.2.6 Validação Cruzada	33
2.3 Redes Neurais Autoassociativas	34
2.4 Métricas para avaliação de desempenho de sistemas de classificação de padrões	37
2.5 Estado da Arte de Redes Neurais Artificiais aplicadas ao problema de diagnóstico de faltas em Transformadores.....	38
3 METODOLOGIA	41
3.1 Introdução.....	41
3.2 Base de dados de transformadores com faltas	41
3.3 Sistema de diagnóstico de faltas via rede Autoassociativas Empilhadas.....	42
3.3.1 Definição das entradas e saídas do sistema	42

3.3.2 Estrutura da rede Autoassociativa empilhada para o diagnóstico de faltas	44
4 RESULTADOS.....	47
4.1 Sistema de diagnóstico baseado em redes Autoassociativas empilhadas	47
4.2 Estado comparativo com método IEC 60599 e Sistemas baseados em Rede Neural	
Perceptron de Múltiplas camadas	50
5 CONCLUSÃO	53
REFERÊNCIAS BIBLIOGRÁFICAS	54
APÊNDICE A	57
APÊNDICE B	66

1 INTRODUÇÃO

O bom funcionamento de um transformador é de grande importância para a integridade de um sistema de energia elétrica, já que este é peça fundamental no tratamento e transmissão da eletricidade e seu estado operacional influencia diretamente na estabilidade do sistema.

Para que sejam evitados desligamentos indevidos, acidentes e prejuízos para o sistema elétrico, é fundamental um plano de manutenção funcional. No que diz respeito à manutenção de equipamentos elétricos, as técnicas tradicionalmente utilizadas são: manutenção preventiva, manutenção corretiva e manutenção preditiva. A manutenção preventiva prega a intervenção no equipamento antes que ele pare de operar, sendo realizada de forma programada e baseada na experiência, estatística ou outro tipo de avaliação, analisando a conveniência ou não da parada do equipamento. Na manutenção corretiva o equipamento opera até ocorrer uma falha que o interrompa ou um defeito que provoque perda parcial ou total da sua função e que justifique a intervenção. A manutenção preditiva prega o monitoramento da tendência e modo de falha de um equipamento, ou seu componente, através do acompanhamento de parâmetros indicativos de sua condição. A principal forma de manutenção adotada é a manutenção preventiva, que promove melhor vida útil dos equipamentos e diminuiu a ocorrência de falhas inesperadas.

O transformador de potência é um dos equipamentos mais caros e importantes de um sistema de potência e, durante sua operação, é submetido a uma variedade de problemas térmicos e elétricos que pode levar a falhas. A detecção e eliminação destas falhas é uma das principais preocupações da manutenção destes equipamentos (CASTRO, 2004).

As falhas nos transformadores geram uma série de gases que são retidos pelo óleo isolante. A concentração e relativa proporção destes gases no óleo é usada na manutenção do equipamento para estimar seu estado operativo. Existem diversos métodos baseados na análise dos gases dissolvidos, ou em inglês *Dissolved Gas Analysis (DGA)*, como Dornenberg, Rogers e IEC 60599 (CASTRO, 2004; MORAIS, 2008). É importante citar que este processo de diagnóstico não é considerado uma ciência, mas uma arte sujeita a variabilidade (N'CHO *et al.*, 2016).

Diversas técnicas de inteligência computacional vêm sendo aplicadas para auxílio ao diagnóstico de faltas em transformadores. Técnicas tais como redes neurais artificiais (GUO; DONG; WU, 2019; KIM *et al.*, 2013; QI *et al.*, 2017), lógica difusa (APTE; SOMALWAR; WAJIRABADKAR, 2018; IRUNGU; AKUMU; MUNDA, 2016; PALKE; KORDE, 2020),

SVM (*Support Vector Machine*) (BENMAHAMED; TEGUAR; BOUBAKEUR, 2018; DHINI *et al.*, 2018; SAHRI; YUSOF, 2015), entre outras.

Apesar dos bons resultados já apresentados na literatura usando técnicas de inteligência computacional para o diagnóstico de transformadores, a pesquisa nesta área ainda tende a evoluir cada vez mais. Alguns pontos ainda vêm sendo tratados por parte dos pesquisadores como por exemplo, a construção de uma base de dados mais completa e consistente para treinamento destes sistemas e a seleção de características que tornem o diagnóstico mais simples e eficiente.

1.1 Objetivo geral

Considerando a importância da análise cromatográfica do óleo (DGA) no planejamento da manutenção dos transformadores de um sistema elétrico, o presente Trabalho de Conclusão de Curso tem por objetivo geral propor um sistema para diagnóstico inteligente do estado do transformador utilizando Redes Neurais Artificiais (Redes Autoassociativas Empilhadas).

Com as Redes Neurais Artificiais é possível a construção de uma metodologia que utiliza uma associação entre os valores das razões entre as concentrações dos gases e o tipo de falha.

Para desenvolvimento do trabalho foi utilizada a Base de Dados TC-10 de equipamentos faltosos, apresentada em (DUVAL; DEPABLA, 2001) e utilizada na publicação do método IEC 60599. Além disso, uma base de dados proveniente de literaturas passadas empregada em (CASTRO, 2004) também foi utilizada.

1.2 Objetivos específicos

Considerando o objetivo principal deste trabalho, foram estabelecidos os seguintes objetivos específicos:

- Buscar, dentre todas as razões de gases possíveis descritas em métodos tradicionais e em novos estudos, aquelas que podem trazer um melhor resultado de diagnóstico;
- Testar a eficiência das redes autoassociativas empilhadas (*Stacked autoencoder*) para o problema de diagnóstico de falhas em transformadores;
- Demonstrar, através de comparação de resultados obtidos, a eficiência do modelo proposto sobre o método IEC 60599 (método criado a partir da base TC 10);

- Desenvolver, a partir da mesma base de dados usada no desenvolvimento do sistema proposto, 2 sistemas de diagnóstico baseados em Redes Perceptron de múltiplas camadas (tipo de rede neural já estabelecido na área, com bons resultados de acurácia apresentados na literatura). Comparar resultados do sistema proposto com as 2 redes desenvolvidas.

1.3 Organização do trabalho

Em adição a este capítulo introdutório, o trabalho é composto por mais quatro capítulos.

O Capítulo 2 apresenta a fundamentação teórica em relação às metodologias de análise de óleo mineral isolante de transformadores, especificamente o método dos gases dissolvidos em óleo, a metodologia que foi empregada no desenvolvimento do sistema inteligente. Além disso, o Capítulo dois apresenta os conhecimentos necessários sobre Redes Neurais Artificiais para a compreensão do sistema inteligente desenvolvido.

O Capítulo 3 apresenta a metodologia utilizada para o desenvolvimento do sistema inteligente, os materiais e métodos utilizados nos testes e para se chegar ao modelo final do sistema proposto.

O Capítulo 4 apresenta os resultados obtidos pelo sistema proposto, bem como por outros sistemas desenvolvidos para comparação, no problema de diagnóstico de faltas em transformadores.

O Capítulo 5 apresenta a conclusão do trabalho e propõe trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Diagnóstico de faltas em transformadores

Os transformadores são os equipamentos mais caros e importantes de um sistema elétrico e estão sujeitos a uma gama variada de falhas elétricas e térmicas, podendo causar problemas como: superaquecimento, arco elétrico e descargas parciais.

A maioria dos transformadores são chamados “à óleo”, o que significa que sua isolação é feita com óleo mineral isolante, que serve como meio isolante para condutores energizados dentro do equipamento. Uma certa quantidade de gases é formada e dissolvida no óleo isolante, tanto durante a operação e degradação normal do equipamento quanto em uma falta elétrica ou térmica, onde a quantidade de gases gerada é muito maior e significativa. O nível de gases dissolvidos no óleo é usado como um indicador primário do estado de saúde do equipamento e, baseada nesta informação, algum tipo de ação de manutenção deve ser tomada (CASTRO, 2004).

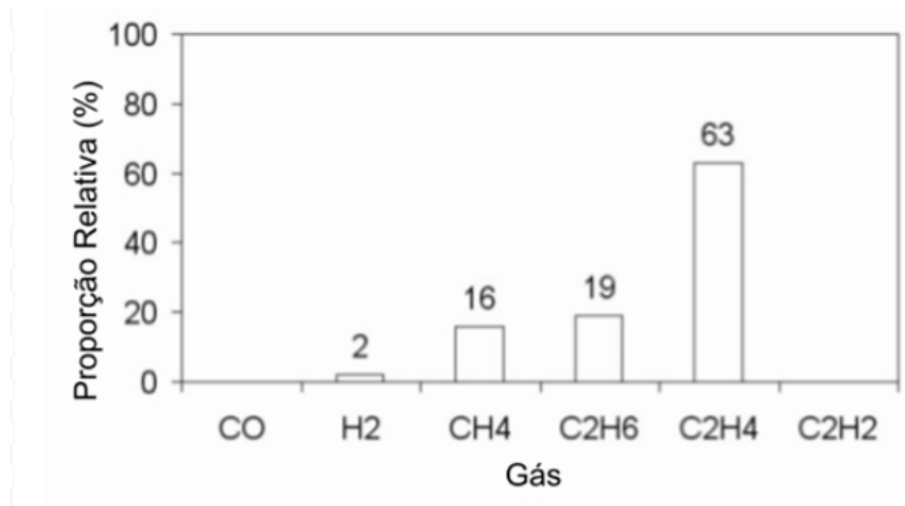
As faltas em transformadores de potência podem ser classificadas nas seguintes categorias: falhas térmicas, descargas parciais e arco elétrico. Tais falhas podem gerar os gases Hidrogênio (H_2), Metano (CH_4), Etano (C_2H_6), Etileno (C_2H_4), Acetileno (C_2H_2), Monóxido de Carbono (CO) e Dióxido de Carbono (CO_2) e com a análise desses gases dissolvidos no óleo é possível não só identificar o tipo de falha, mas também a severidade da mesma, possibilitando a elaboração de um plano de manutenção dos equipamentos (CASTRO, 2004). Tais falhas e exemplos de suas possíveis causas podem ser vistas na Tabela 1 a seguir:

Tabela 1 – Falhas típicas em transformadores de potência

Tipo de falha	Exemplos de causas
Descarga parcial	Descargas em cavidades preenchidas com gás, resultante de impregnação incompleta, alta umidade de papel, óleo super saturado.
Descargas de baixa energia	Faiscamento ou arco entre más conexões de diferentes potenciais ou potenciais flutuantes de anéis de proteção, toróides, discos adjacentes ou condutores de bobinas, bresagem quebrada ou loops fechados no núcleo. Descargas entre peças de fixação, buchas e tanque, alta tensão e terra no interior do enrolamento, nas paredes do tanque. Rastreamento em blocos de madeira, cola de viga isolante, espaçadores de enrolamento. Decomposição do óleo, corrente de ruptura do seletor.
Descargas de alta energia	<i>Flashover</i> , rastreamento ou arco de alta energia local ou com seguimento de energia. Curto circuito entre baixa tensão e terra, conectores, bobinamento e tanque, barra de cobre e tanque, enrolamento e núcleo, duto de óleo, torre. Loops fechados entre dois condutores adjacentes ao redor do fluxo magnético principal, parafusos isolados de núcleo, anéis metálicos que seguram as pernas do núcleo.
Falha térmica $t < 300^{\circ} C$	Sobrecarga de transformadores em situações de emergência. Itens bloqueadores do fluxo de óleo nos enrolamentos.
Falha térmica $300^{\circ} C < t < 700^{\circ} C$	Contatos defeituosos entre ligações aparafusadas (particularmente entre barras de alumínio), contatos deslizantes dentro do interruptor seletor (formação pirolítica de carbono), ligações de cabos e tirantes de buchas. Correntes circulantes entre braçadeiras e parafusos do jugo, braçadeiras e laminação, em cabos de terra, soldaduras defeituosas ou braçadeiras em escudos magnéticos. Isolamento com abrasão entre condutores paralelos adjacentes em enrolamentos.
Falha térmica $t > 700^{\circ} C$	Grande circulação de corrente no tanque e no núcleo. Pequenas correntes na parede do tanque criadas por um grande campo magnético descompassado.

As falhas térmicas ocorrem entre as temperaturas de 150°C e 500°C, podendo atingir valores acima. No caso do superaquecimento de óleo a falha produz, principalmente, Etileno e Metano, com pequenas quantidades de Hidrogênio e Etano, como exemplificado na Figura 1 (IEEE, 2009).

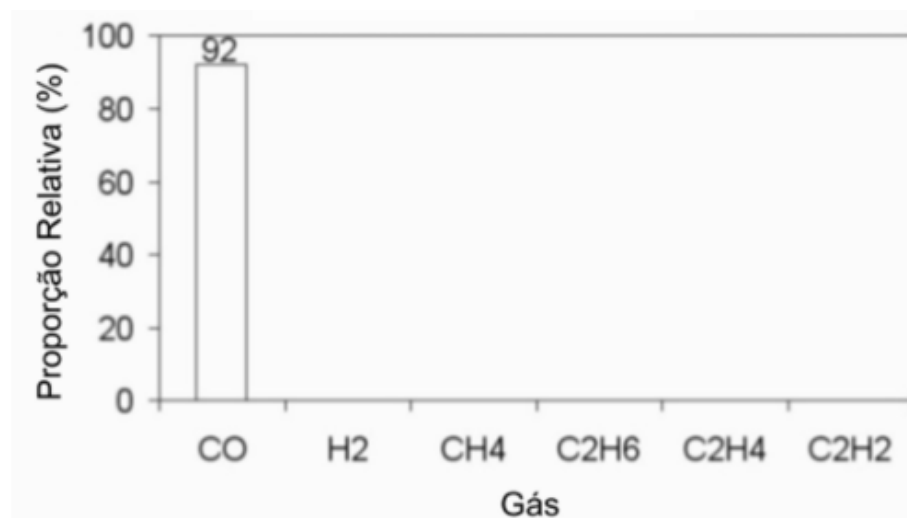
Figura 1 – Superaquecimento de Óleo



Fonte: (IEEE, 2009)

A falha térmica de superaquecimento da celulose provoca grandes quantidades de dióxido e monóxido de carbono (IEEE, 2009), como apresentado na Figura 2.

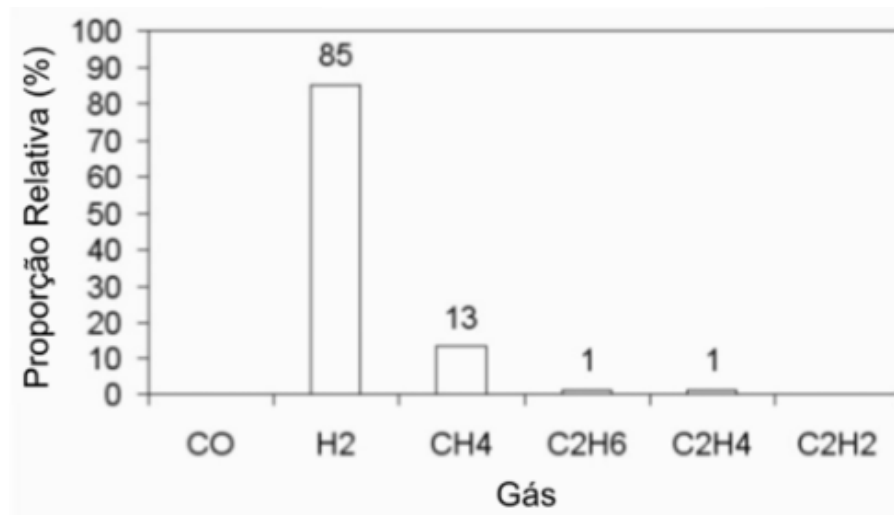
Figura 2 – Superaquecimento de Celulose



Fonte: (IEEE, 2009)

A falha elétrica de descargas parciais de baixa intensidade produzem, principalmente, Hidrogênio, com menores quantidades de Metano e pouca quantidade de Acetileno (IEEE, 2009). Na medida que a intensidade aumenta, as quantidades de Acetileno e Etileno aumentam significativamente. A Figura 3 ilustra este caso.

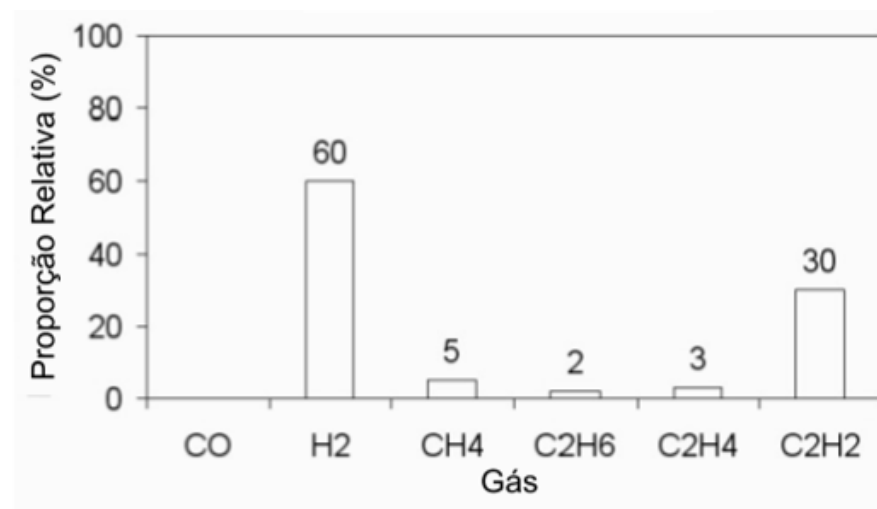
Figura 3 - Descarga Parcial de baixa intensidade em óleo



Fonte: (IEEE, 2009)

A falha elétrica de alta intensidade, arco elétrico, é mais preocupante. Grande quantidade de Hidrogênio e Acetileno são produzidos (Figura 4), com pequenas quantidades de Metano e Etileno (IEEE, 2009).

Figura 4 - Arco elétrico



Fonte: (IEEE, 2009)

Baseados na quantidade e proporção dos gases presentes em óleo, vários métodos foram criados para realizar o diagnóstico do estado de saúde de um equipamento. O método da análise

dos gases dissolvidos em óleo (*Dissolved Gas Analysis - DGA*) é um método confiável de diagnóstico para as faltas no equipamento (IEEE, 2009).

Todavia, é importante salientar que a análise DGA não é uma tarefa fácil, já que, por muitas vezes, o método DGA pode não oferecer um diagnóstico preciso e até um errôneo. Segundo (IEEE, 2009), a análise DGA e o processo de interpretação dos gases não é uma ciência, por estar sujeita a variações, principalmente por conta de diferenças estruturais entre equipamentos, fabricação, histórico, condições ambientais, etc. Sendo a falta de correlação e dados disponíveis que relacionem os gases com as falhas realmente encontradas em um equipamento os principais obstáculos para o procedimento de identificação de faltas como uma ciência exata (N'CHO *et al.*, 2016).

Devido a tais ambiguidades, vários métodos de análise DGA são usados de maneira complementar. A seguir, serão apresentados os principais métodos de análise DGA.

2.1.1 Método de Dornenberg

Pioneiro na aplicação de metodologias DGA no diagnóstico de falhas em Transformadores (CASTRO, 2004), o método de Dornenberg é usado até hoje e se baseia em 4 razões das concentrações de gases (R1, R2, R3 e R4) como mostrado na Tabela 2.

Tabela 2 - Razões de Dornenberg

Relações entre concentrações de gases				
Tipo de Falha	Relações principais		Relações auxiliares	
	CH_4/H_2	C_2H_2/C_2H_4	C_2H_6/C_2H_2	C_2H_2/CH_4
Ponto Quente	>1	<0.75	>0.4	<0.3
Descarga parcial	<0.1	Não signif..	>0.4	<0.3
Outros tipos de descarga	<1 e >0.1	>0.75	<0.4	>0.3

O estabelecimento para as condições de validade de aplicação do método é feito em função da concentração individual dos gases presentes nas relações (MORAIS, 2008), como pode ser visto na Tabela 3.

Tabela 3 – Concentração dos gases para validação do método de Dornenberg

Tipos de gás	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Concentrações ppm (v/v)	200	50	15	60	15

O critério indica que, para ser aplicável, no mínimo um dos gases que compõe as relações principais tenha uma concentração superior ao dobro do valor na Tabela 3 e que para os gases auxiliares pelo menos um dos gases tenha uma concentração maior que o especificado na Tabela 3.

2.1.2 Método de Rogers

O método de Rogers utiliza 4 razões e 5 gases para se chegar a um diagnóstico, como apresentado na Tabela 4. Tem-se um código que é combinado e a partir dessa combinação chega-se no diagnóstico final como mostrado na Tabela 5.

Tabela 4 - Razões e códigos

Relação de Gases	Faixa de Variação	Código
CH_4/H_2	≤ 0.1	5
	$> 0.1 < 1$	0
	$\geq 1 < 3$	1
	≥ 3	2
C_2H_6/CH_4	< 1	0
	≥ 1	1
C_2H_4/C_2H_6	< 1	0
	$\geq 1 < 3$	1
	≥ 3	2
C_2H_2/C_2H_4	< 0.5	0
	$\geq 0.5 < 3$	1
	≥ 3	2

Tabela 5 – Combinações de códigos e diagnóstico

CH_4/H_2	C_2H_6/CH_4	C_2H_4/C_2H_6	C_2H_2/C_2H_4	Diagnóstico
0	0	0	0	Deterioração Normal
5	0	0	0	Descargas Parciais
1/2	0	0	0	Sobreaquecimento abaixo de 150°C
1/2	1	0	0	Sobreaquecimento de 150°C – 200°C
0	1	0	0	Sobreaquecimento de 200°C – 300°C
0	0	1	0	Sobreaquecimento de condutores
1	0	1	0	Correntes de circulação nos enrolamentos
1	0	2	0	Correntes de circulação no núcleo e tanque, sobreaquecimento em conexões
0	0	0	1	Descarga contínua
0	0	1/2	1/2	Arco com alta energia
0	0	2	2	Descarga contínua com baixa potência
5	0	0	1/2	Descarga parcial envolvendo o papel

Foi um dos primeiros métodos a ser aplicado e aceito devido à boa acurácia (CASTRO, 2004), todavia, assim como Dornenberg, pode chegar em combinações que não estão previstas nos códigos de diagnóstico.

2.1.3 Método IEC 60599

O método de diagnóstico proposto na norma IEC 60599 (IEC 60599, 1999), é ilustrado na Tabela 6. Se ao menos um gás tem valor de concentração acima do normal, então as razões entre os gases chave são calculadas e o método é aplicado (CASTRO, 2004).

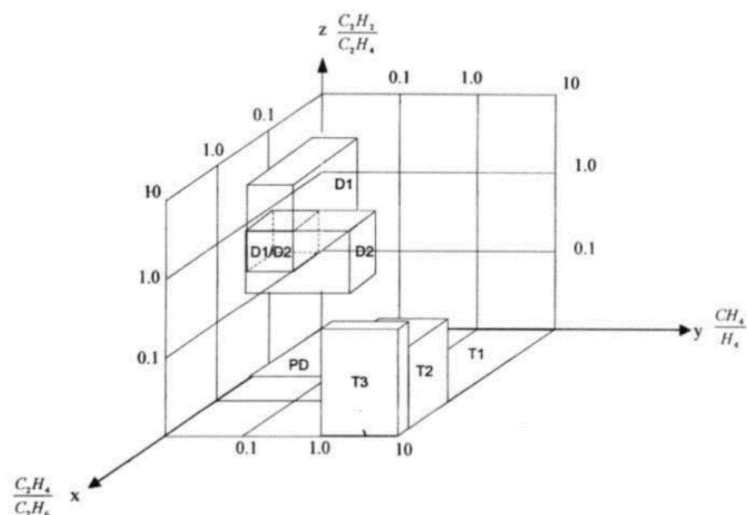
Tabela 6 - Método IEC

C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6	Falta característica
Não significativo	<0.1	<0.2	Descargas parciais
>1	0.1 - 0.5	>1	Descargas de baixa energia
0.6 - 2.5	0.1 - 1	>2	Descargas de alta energia
Não significativo	>1 mas não significativo	<1	Falha térmica $T < 300^{\circ}C$
<0.1	>1	1 - 4	Falha térmica $300^{\circ}C < T < 700^{\circ}C$
<0.2	>1	>4	Falha térmica $T > 700^{\circ}C$

O código da Tabela 6 é resultado de uma revisão no banco de dados do IEC TC 10, da publicação anterior, IEC 599 (DUVAL; DEPABLA, 2001). IEC TC 10 é um banco de dados que agrega resultados de concentrações de gás chave relacionados com falhas em equipamentos inspecionadas visualmente por especialistas em campo.

Este método é uma variação do método de Rogers, excluindo-se a concentração C_2H_6/CH_4 . Deve-se notar que uma das grandes desvantagens do método IEC é o diagnóstico “não-decisão”, sendo incapaz de diagnosticar certo tipo de falha para determinadas razões. Da Tabela 5 podemos ver que o método também tem o problema de resultados ambíguos para uma mesma amostra entre as faltas do tipo descarga de alta e baixa potência, atribuindo, por vezes, ambos os diagnósticos (CASTRO, 2004; FAIZ; SOLEIMANI, 2017). A Figura 5 apresenta, visualmente, os problemas de “sem resultado” e resultados ambíguos da tabela IEC 60599.

Figura 5 – Representação gráfica do código IEC 60599

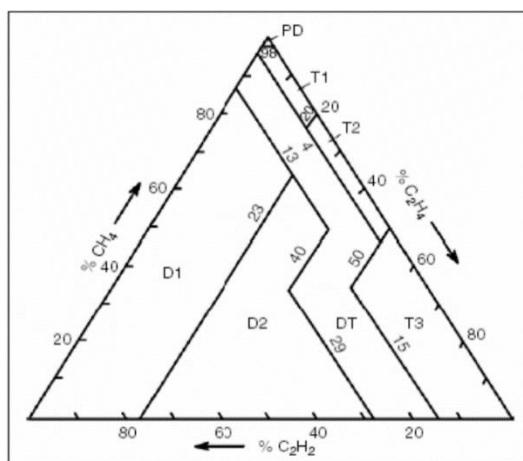


Fonte: (CASTRO, 2004)

2.1.4 Método Duval

O método proposto por Duval leva em conta somente a concentração percentual relativa dos gases acetileno, metano e etano (MORAIS, 2008). A Figura 6 demonstra o chamado "Triângulo de Duval", uma representação gráfica visual das áreas de um triângulo e um ponto, cujas coordenadas se dão de acordo com a proporção dos gases. A localização do ponto será o diagnóstico.

Figura 6 - Triângulo de Duval



Fonte: (CASTRO, 2004)

2.1.5 Método do Gás Chave

Este método identifica o gás chave para cada tipo de falha e usa seu percentual para fazer a classificação da falha. A Tabela 7 ilustra o processo, na qual percentual do gás é baseado no total de gases combustíveis dissolvidos (TDCG) e é um número aproximado (MORAIS, 2008).

Tabela 7 - Método Gás Chave

Falta	Gás chave	Critério	Percentual do gás
Arco	Acetileno (C_2H_2)	Grande quantidade de H_2 e C_2H_2 e menor quantidade de CH_4 e C_2H_4 . CO e CO_2 podem ser encontrados se a celulose estiver envolvida	H_2 : 60% C_2H_2 : 30%
Corona	Hidrogênio (H_2)	Grande quantidade de H_2 , algum CH_4 , com pequena quantidade de C_2H_6 e C_2H_4 . CO e CO_2 podem ser comparáveis se a celulose estiver envolvida	H_2 : 85% CH_4 : 13%
Sobreaquecimento do óleo	Etileno (C_2H_4)	Grande quantidade de C_2H_4 , menor quantidade de C_2H_6 , alguma quantidade de CH_4 e H_2 . Traços de CO	C_2H_4 : 63% C_2H_6 : 20%
Sobreaquecimento da celulose	Monóxido de Carbono CO	Grande quantidade de CO e CO_2 . Gases hidrocarbonetos podem existir	CO : 92%

2.2 Redes Neurais Artificiais

Redes Neurais Artificiais (RNAs) são técnicas de inteligência computacional que se baseiam no funcionamento do cérebro humano, um órgão o qual acredita-se ter aproximadamente 10 bilhões de unidades de processamento interconectadas, os neurônios, sendo cada um desses possuidores da capacidade de transmitir e receber energia. O cérebro é capaz de realizar tarefas complexas como reconhecimento de padrões, percepção e controle

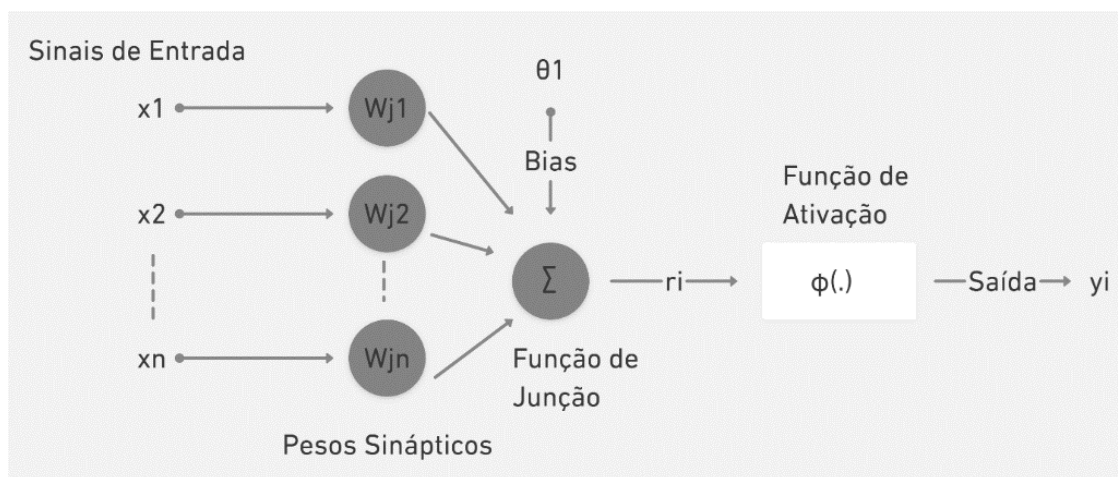
motor devido à capacidade de se auto-organizar e formar ligações de neurônios, as ligações sinápticas (CASTRO, 2004; MORAIS, 2008).

A Rede Neural Artificial possui um grande número de unidades de processamento interconectados, os chamados neurônios artificiais, interligados por um grande número de interconexões (sinapses artificiais), representados por vetores e matrizes de pesos sinápticos (SILVA, 2016). A computação da RNA é performada por estes neurônios que operam em paralelo e a representação do conhecimento adquirido durante o processo de aprendizagem se dá mediante distribuição e reajuste dos pesos sinápticos entre esses neurônios. Durante o treinamento, padrões de treinamento (dados do problema) são apresentados à rede neural continuamente com o objetivo de minimizar uma função custo através da modificação dos pesos sinápticos (CASTRO, 2004).

2.2.1 Neurônio Artificial

A unidade de processamento da RNA, o neurônio artificial, tem seu funcionamento exemplificado na Figura 7 e foi inspirado a partir da análise de geração e propagação de impulsos elétricos pela membrana celular dos neurônios (SILVA, 2016). O neurônio artificial é composto por três elementos funcionais: os pesos sinápticos, a função de junção (soma) e a função de ativação.

Figura 7 – Neurônio Artificial



Fonte: Autor

Os neurônios artificiais possuem terminais de entrada e saída e podem ser lineares ou não lineares (dependendo da função de ativação), fornecendo saídas tipicamente contínuas. Os neurônios realizam a função de coletar os sinais existentes em suas entradas, ponderá-los pelos pesos, agregá-los de acordo com a função de junção (*soma*) e produzir uma resposta, levando em consideração sua função de ativação inerente (SILVA, 2016).

Os pesos regulam a conexão entre a fonte de sinal e os neurônios, atenuando ou amplificando, por um fator de peso w_{ij} . Diz-se que o fator de peso é positivo se a respectiva sinapse, advinda da fonte externa ou dos outros neurônios é excitatória, e negativo se a sinapse é inibitória. Portanto, quanto maior o valor do peso, mais intensa é a ligação entre os neurônios (CASTRO, 2004).

A função de junção implementa a soma ponderada das entradas de acordo com a seguinte expressão:

$$r_j = \sum_{i=1}^p w_{ij}x_i + \theta_j \quad (1)$$

Onde θ_j é um bias aplicado externamente com a função de aumentar ou diminuir o sinal da soma.

A função de ativação recebe o sinal da função junção e calcula a estimulação produzida por aquele sinal. O neurônio pode então, baseado nesta estimulação, produzir uma saída, que pode ser linear ou não linear e se dá mediante a seguinte expressão:

$$y_j = \varphi\left(\sum_{i=1}^p w_{ij}x_i + \theta_j\right) \quad (2)$$

Onde φ é a função de ativação, também chamada de função de transferência.

2.2.2 Funções de ativação

A função de ativação do neurônio tem por objetivo limitar a saída, ou seja, tomar uma decisão sobre o que fazer com o valor proveniente da função de junção. Existem diversos tipos de função de ativação, sendo as mais utilizadas: degrau, linear, tangente hiperbólica e sigmoide (MORAIS, 2008).

2.2.2.1 Função Degrau

O resultado produzido pela aplicação desta função assumirá valores unitários positivos quando o potencial de ativação do neurônio for maior ou igual a zero; e valores nulos para outros casos. Esta função é dada por:

$$\varphi(r_j) = \begin{cases} 1 & \text{se } r_j \geq 0; \\ 0 & \text{se } r_j < 0; \end{cases} \quad (3)$$

2.2.2.2 Função Linear

Também chamada de função de identidade, produz resultados de saída idênticos aos valores do potencial de ativação. Esta função é dada por:

$$\varphi(r_j) = r_j \quad (4)$$

2.2.2.3 Função Sigmoidal

Também chamada de função logística, esta é uma função amplamente utilizada para redes neurais. É uma função estritamente crescente que apresenta propriedades assintóticas e de suavidade sendo que o resultado da sua aplicação sempre assumirá valores reais entre zero e um. Esta função é dada por:

$$\varphi(r_j) = \frac{1}{1 + e^{-ar_j}} \quad (5)$$

Onde a é uma constante real associada ao nível de inclinação da função.

2.2.2.4 Função Tangente Hiperbólica

Ao aplicar esta função, diferente da função sigmoidal, seu resultado sempre assumirá valores reais entre -1 e 1. Esta função é dada por:

$$\varphi(r_j) = \frac{1 - e^{-ar_j}}{1 + e^{-ar_j}} \quad (6)$$

Onde a também é uma constante real associada ao nível de inclinação da função.

2.2.3 Arquiteturas de RNA

Segundo (SILVA, 2016), a arquitetura de uma rede neural é a maneira pela qual os neurônios estão arranjados, ou dispostos, uns em relação aos outros. Uma rede neural pode ser dividida em três partes, chamadas de camadas, da seguinte forma:

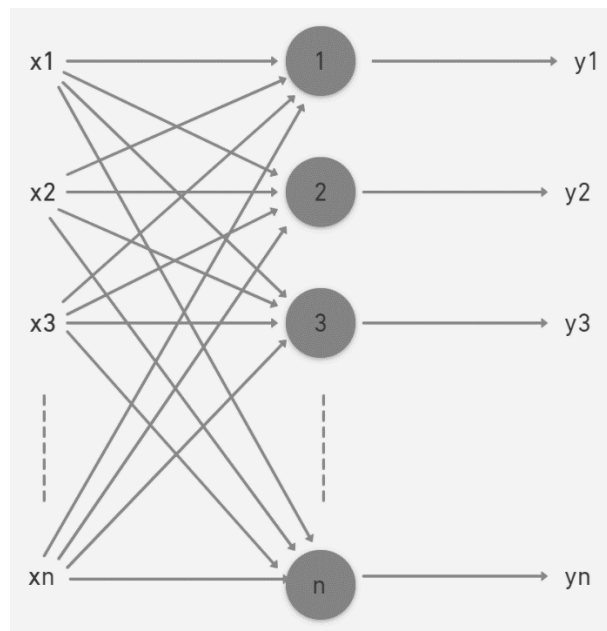
- Camada de entrada: Responsável pelo recebimento dos sinais do meio externo;
- Camada escondida: Compostas por neurônios que possuem a responsabilidade de extrair as características dos sinais e executar o processamento interno da rede;
- Camada de saída: Também constituída por neurônios, responsável pela produção e apresentação do resultado final da rede.

Considerando a disposição e forma de interligação dos neurônios, as principais arquiteturas de redes neurais são: redes alimentadas adiante (*Feedforward*) de única camada, redes alimentadas adiante com múltiplas camadas e redes com realimentação (SILVA, 2016).

2.2.3.1 Redes alimentadas adiante de uma única camada

As redes de uma única camada possuem uma camada de entrada e uma única camada de neurônios, que é a própria camada de saída. O fluxo de informações segue da camada de entrada (nós fonte) para a saída (nós computacionais). A camada de entrada não é considerada na contagem do número de camadas já que é considerada especial e seus neurônios têm o dever somente de distribuir cada entrada aos nós de computação seguintes e a quantidade de saídas sempre coincidirá com o número de neurônios (MORAIS, 2008). Essas redes têm emprego típico em problemas de classificação de padrões e filtragem linear, sendo a rede *Perceptron* um dos exemplares mais famosos e utilizados dessa arquitetura (SILVA, 2016). A Figura 8 ilustra esta arquitetura.

Figura 8 – Rede de uma única camada

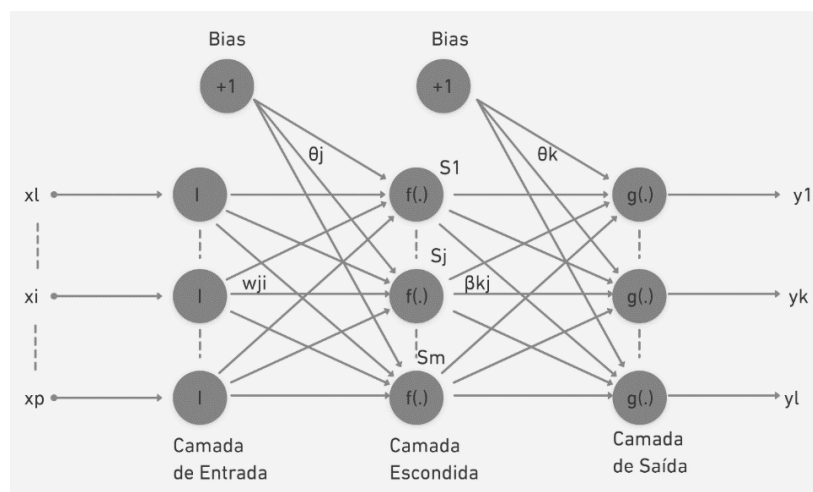


Fonte: Autor

2.2.3.2 Redes alimentadas adiante com Múltiplas Camadas

As redes de camadas múltiplas são constituídas pela presença de uma ou mais camadas escondidas de neurônios, como mostra a Figura 9. Tendo como exemplo a rede Perceptron de múltiplas camadas (MLP- *Multi layer Perceptron*), elas são, hoje em dia, uma das redes mais utilizadas em diversas aplicações como aproximação de funções, classificação de padrões, identificação de sistemas, otimização, controle de processos, etc (MORAIS, 2008; SILVA, 2016).

Figura 9 – Rede com múltiplas camadas



Fonte: Autor

Uma MLP é um modelo consistente de sucessivas camadas de neurônios. Os neurônios de cada camada estão conectados com as outras mediante os pesos sinápticos. A informação flui de uma camada para a sua subsequente, começando pela camada de entrada que recebe os sinais advindos do meio externo, uma ou mais camadas escondidas e a camada de saída (CASTRO, 2004; SILVA, 2016).

Considerando a MLP da Figura 9 com apenas uma camada escondida, cada neurônio da camada oculta calcula:

$$sj = f \sum_{i=1}^p x_i w_{ij} + \theta_j \quad (7)$$

Onde x_i é a i -ésima entrada, w_{ij} é o peso entre o neurônio de entrada i e o neurônio escondido j , θ_j é o bias do j -ésimo neurônio escondido e $f(.)$ é a função de ativação do neurônio escondido.

Considerando a camada escondida, cada neurônio calcula:

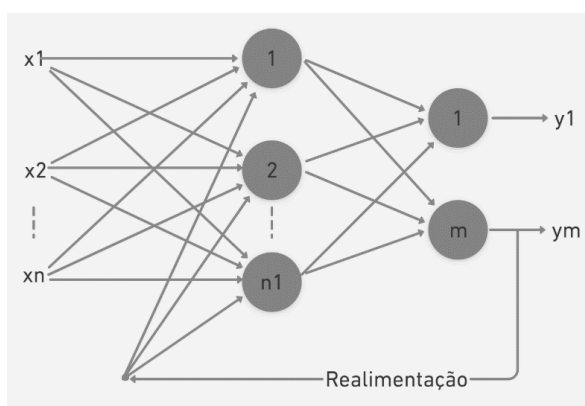
$$y_k = g \left(\sum_{j=1}^m \beta_{kj} S_j \right) + \theta_k \quad (8)$$

Onde β_{kj} é o peso entre o neurônio escondido j e o neurônio de saída k , y_k é a k -ésima saída da rede, θ_k é o bias da k -ésima saída e $g(.)$ é a função de ativação neurônio de saída.

2.2.3.3 Redes com realimentação

Este tipo de rede difere dos tipos *feedforward* pois nas saídas dos neurônios, elas são realimentadas como sinais de entrada para outros neurônios, ou seja, a saída do neurônio é retroalimentada na entrada dos outros neurônios das camadas anteriores. A realimentação as qualifica para o processamento dinâmico de informações, isto é, podem ser usadas em sistemas variantes em relação ao tempo, como previsão de séries temporais, otimização e identificação de sistemas, controle de processos, etc (SILVA, 2016). A Figura 10 ilustra uma rede realimentada.

Figura 10 – Rede com realimentação



Fonte: Autor

2.2.4 Paradigmas de Aprendizagem

Uma das propriedades de maior destaque em uma rede neural é a sua capacidade de aprender a partir da apresentação de amostras que representem o comportamento dos sistemas e de melhorar seu aprendizado continuamente, generalizando soluções e produzindo uma saída aproximada daquela desejada a partir de quaisquer sinais inseridos em sua entrada (MORAIS, 2008; SILVA, 2016).

O aprendizado da rede se dá mediante um processo iterativo de ajustes aplicados nos pesos sinápticos da rede e, segundo (SILVA, 2016), o processo de treinamento de uma rede neural: “Consiste de uma aplicação de passos ordenados que sejam necessários para sintonização dos pesos sinápticos e limiares de seus neurônios, tendo-se como objetivo final a generalização de soluções a serem produzidas pelas suas saídas, cujas respostas são representativas do sistema físico em que estão mapeando”.

Para realizar o processo de aprendizagem, é importante ter um modelo do ambiente de operação da rede neural, isto é, precisa-se saber quais informações estão disponíveis para a rede neural. Isso se dá em um processo denominado paradigma de aprendizagem, que se subdivide em duas principais formas: aprendizagem supervisionada e não supervisionada.

A aprendizagem supervisionada tem por estratégia ter disponível, para cada amostra de treinamento, uma respectiva saída desejada que é previamente conhecida. A saída da rede neural é comparada com a desejada e é calculado um erro. Os pesos são ajustados de maneira a minimizar este erro (CASTRO, 2004).

A aprendizagem não supervisionada não tem as saídas desejadas, os pesos são determinados como resultado de um processo auto-organizado, ou seja, as conexões se auto-organizam em relação às particularidades existentes entre elementos existentes nas amostras, identificando subconjuntos (*clusters*) que contenham similaridades (CASTRO, 2004).

Dentre os algoritmos usados na aprendizagem supervisionada está o *Backpropagation* como o mais utilizado para o treinamento de uma MLP (CASTRO, 2004; MORAIS, 2008).

2.2.4.1 Algoritmo Backpropagation

O algoritmo *Backpropagation* é um processo do tipo iterativo baseado na minimização do erro médio quadrático que usa o método chamado de gradiente descendente para atualização dos pesos. Também chamado de regra Delta generalizada, é realizado com a aplicação de duas fases bem específicas:

- Propagação adiante (*forward*): As informações inseridas na entrada da rede são propagadas camada a camada até a produção da sua respectiva saída. Então o erro entre a saída obtida e desejada é calculado;
- Propagação reversa (*backward*): As alterações dos pesos e limiares da rede são executadas nessa fase. O erro, já calculado, agora é retropropagado por cada camada para que, com esta referência de valor, os pesos sejam atualizados.

Após atualizar todos os pesos da rede o processo recomeça. Um ciclo completo é chamado de época e quando esta é completa a tendência é que o erro quadrático médio diminua, até alcançar um mínimo que pode ser pré-estabelecido para terminar o treinamento.

O algoritmo *Backpropagation* é eficiente para o treinamento de redes de múltiplas camadas, pois atualiza seus pesos sinápticos através da minimização do erro quadrático, que é possível pela utilização do método do gradiente descendente (MORAIS, 2008).

Considerando a rede neural da Figura 9, o sinal de erro do neurônio de saída k , na iteração n é calculado por:

$$e_k(n) = d_k(n) - y_k(n) \quad (9)$$

Para todos os neurônios de saída no conjunto C, a soma instantânea do erro quadrático é dada por:

$$E(n) = \sum_{k \in C} e_k^2(n) \quad (10)$$

E para todos os padrões N apresentados à rede, o erro médio quadrático é definido por:

$$E_{medio} = \frac{1}{N} \sum_{n=1}^N E(n) \quad (11)$$

O erro quadrático médio E_{medio} se dá em função dos parâmetros livres da rede neural (pesos e bias) e ele representa a função custo que precisa ser minimizada durante o processo (CASTRO, 2004).

O ajuste dos pesos, na camada escondida l , é performedo de acordo com a regra delta generalizada:

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \eta \delta_j^{(l)}(n) y_i^{(i-1)}(n) \quad (12)$$

Onde η é a taxa de aprendizado e $\delta_j^{(l)}(n)$ é o gradiente local, calculado por:

$$\delta_j^{(l)}(n) = \begin{cases} e_j^{(L)}(n) g_j(r_j^{(L)}(n)), & \text{neurônio } j \text{ na camada de saída } L \\ f'_j(r_j^{(l)}(n)) \sum \delta_j^{(l+1)}(n) w_{kj}^{(l+1)}(n), & \text{neurônio } j \text{ na camada escondida } l \end{cases} \quad (13)$$

É interessante destacar que, como pode ser visto na equação (13), a maior restrição para uma função de ativação é que ela tem de ser diferenciável (CASTRO, 2004).

2.2.5 Generalização em RNA

Generalização é a capacidade de uma RNA responder corretamente à padrões não apresentados para a rede durante a fase de treino. Estes padrões devem pertencer a mesma distribuição de dados da base usada para treino.

Considerada uma habilidade fundamental de uma RNA, alguns fatores devem ser levados em conta para se obter uma alta capacidade de generalização, sendo influenciada pelos seguintes fatores:

- Tamanho e representatividade do conjunto de dados de treinamento;
- Arquitetura da rede neural;
- Complexidade física do problema abordado.

Não há uma regra para se obter o tamanho do conjunto de treinamento, cada problema requer uma quantidade de amostras capaz de representá-lo (MORAIS, 2008).

A escolha da arquitetura certa é um dos maiores desafios na capacidade de generalização. Quando a complexidade do modelo é maior que a necessária, a rede fica superajustada (*overfitting*) aos dados de treinamento, respondendo de maneira inadequada a outros dados no processo de treinamento e teste, reduzindo a capacidade de generalização. Porém, se a complexidade do problema supera a da rede, esta não será capaz de representar o domínio do problema, caracterizando assim o fenômeno de *underfitting* (CASTRO, 2004; MORAIS, 2008).

2.2.6 Validação Cruzada

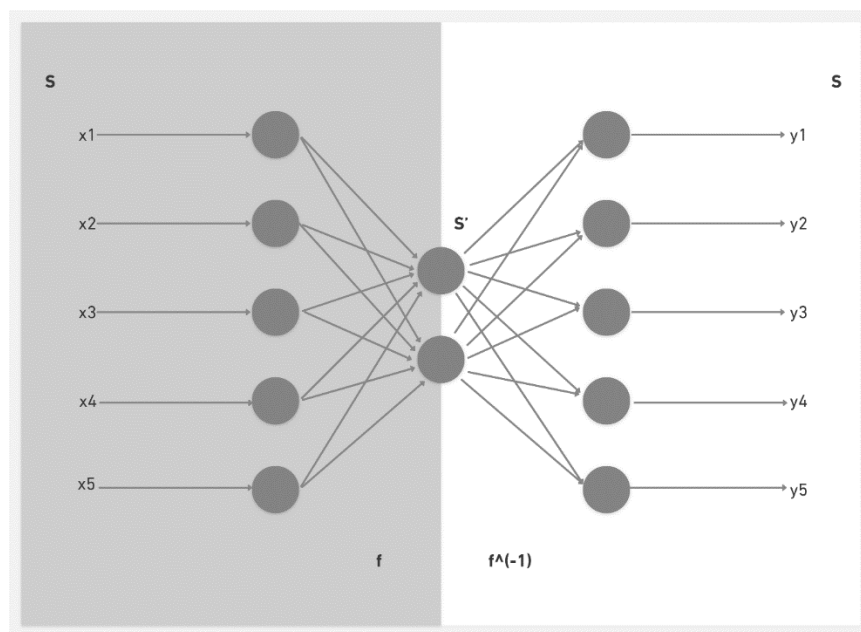
Na aprendizagem por retropropagação do erro, espera-se que a rede neural aprenda o suficiente de modo a ser capaz de generalizar no futuro. Assim, o processo de aprendizagem se torna um processo de definição de parâmetros, estruturas e modelos de uma rede neural (MORAIS, 2008).

Para que as escolhas sejam acertadas para uma boa generalização, o treinamento pode ser realizado usando o processo de validação cruzada, cuja motivação é validar o modelo com um conjunto de dados diferentes daquele usado para estimar os parâmetros. O conjunto de dados é dividido em três: treinamento, validação e teste. O conjunto de teste não é utilizado até que o treinamento finalize. O conjunto de validação é utilizado durante fase de treino para validar o modelo. Caso o erro de validação cresça enquanto o de treino decresce, o treino deve ser finalizado. Desta forma, utiliza-se a validação cruzada para descartar arquiteturas que não apresentem boa capacidade de generalização. Para definição final da arquitetura a ser utilizada, aplica-se a base de teste para verificação da capacidade de generalização em dados nunca apresentados a rede e desta forma garantir a escolha da melhor topologia.

2.3 Redes Neurais Autoassociativas

Uma Rede Neural Autoassociativa (em inglês, *Auto Encoder* – AE) é um caso especial de uma rede Perceptron de múltiplas camadas (MLP), na qual o número de neurônios da camada de entrada é igual ao número de neurônios da camada de saída. Este tipo de rede neural é treinada para mapear a função identidade entre a entrada e saída, ou seja, os padrões de entrada devem ser reconstruídos na saída da rede (MORAIS, 2019). A Figura 11 apresenta a estrutura de uma AE

Figura 11 – Estrutura de uma rede Autoassociativa



Fonte: Autor

A Rede AE mapeia as entradas do espaço S (meio externo) para o espaço S' (interno), aproximando uma função f , e obtém o mapeamento inverso de S' para S (saída da rede) aproximando a função inversa f^{-1} sendo esse processo chamado de codificação e decodificação da informação. Se o vetor de saída for igual ao vetor de entrada, todas as informações fluem através do gargalo de S e são recompostas em S' . Assim, a RNA AE aprende os dados reais e os armazena nos seus pesos (MORAIS, 2019).

A camada intermediária da rede AE é responsável por extrair características das informações fornecidas pelos neurônios de entrada. É esta camada que permite a rede armazenar somente o conhecimento relacionados às principais características do conjunto de dados do problema. Assim, a camada escondida é responsável por codificar os dados enquanto a camada

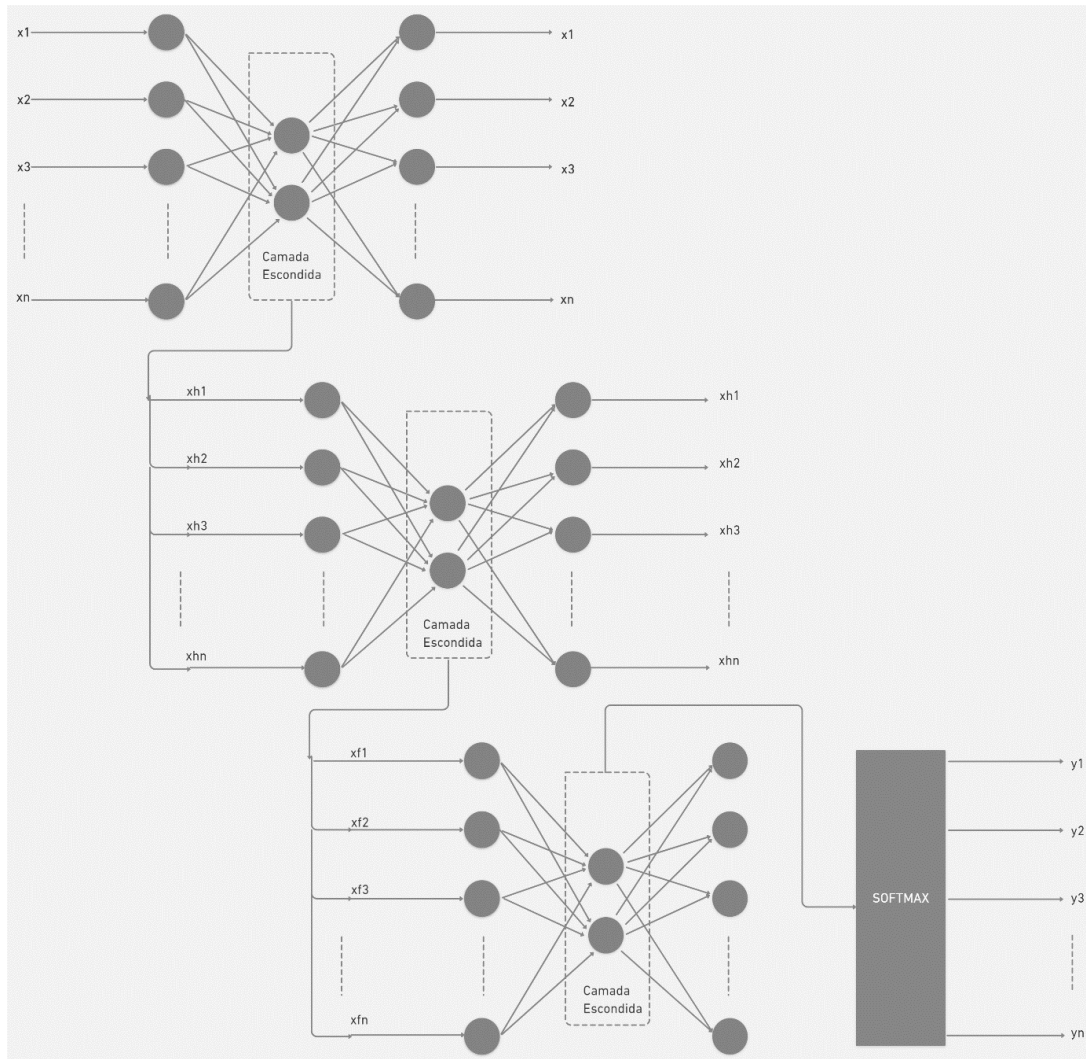
de saída é a responsável por reconstruir as entradas originais. Desta forma, a rede realiza a aproximação da função identidade (MORAIS, 2019).

Para o treinamento de uma rede AE, utiliza-se o algoritmo *Backpropagation* padrão, já que ela é um caso especial de uma rede MLP, com minimização do erro médio quadrático entre os valores dos neurônios de entrada e saída. O treinamento é feito de maneira não supervisionada, já que o objetivo da RNA AE é minimizar o erro entre sua amostra na entrada e a reconstrução na saída, ao invés de minimizar o erro entre a saída e um rótulo externo (MORAIS, 2019).

Para problemas de classificação, é bastante utilizado a arquitetura chamada de “Autoassociativas Empilhadas” ou, em inglês, *Stacked Auto Encoder* (SAE) com uma camada de saída onde a função *Softmax* é aplicada. A estrutura pode ser vista da Figura 12. Na SAE, as redes autoassociativas se conectam, onde o sinal de saída da camada escondida de uma AE (características extraídas a partir da entrada aplicada nesta AE) serve como sinal de entrada para outra AE. Por último, tem-se a camada responsável pela classificação, a função *Softmax* que recebe como entrada o sinal extraído pela camada escondida da última AE empilhada.

Com o uso das redes autoassociativas empilhadas pretende-se a extração de características da base de dados original, de tal forma que estas possam ser utilizadas para o problema de classificação através do uso da função *softmax* na saída da rede.

Figura 12 – Stacked Autoencoder



Fonte: Autor

A função *Softmax* transforma as saídas correspondentes para cada classe em valores entre zero e um, os dividindo pela soma das saídas. Isso, essencialmente, gera a probabilidade de a entrada estar em uma determinada classe. Esta função tem por definição:

$$S(x^t w_j) = \frac{e^{x^t w_j}}{\sum_{k=1}^K e^{x^t w_k}} \quad (14)$$

Onde j corresponde ao índice da classe em que se deseja avaliar, w representa a matriz de pesos sinápticos de uma amostra, x é o vetor de entrada e k indica os índices pertencentes as outras classes avaliadas no modelo (MORAIS, 2019).

2.4 Métricas para avaliação de desempenho de sistemas de classificação de padrões

Para a avaliação de uma rede neural para problemas de classificação de padrões, geralmente se utilizam algumas principais métricas, sendo elas: matriz de confusão, acurácia, precisão, sensibilidade e F_score.

A matriz de confusão é uma ferramenta gráfica importante pois permite analisar rapidamente o desempenho do sistema, os valores que compõem a matriz advêm da comparação entre o conjunto de teste e a predição do sistema, que computa todos os acertos e erros. Ela nos mostra a frequência de cada um dos tipos abaixo:

TP significa verdadeiro positivo, no caso de diagnóstico de faltas indicaria que a amostra foi classificada corretamente com tendo uma determinada falta.

TN significa verdadeiros negativo, no caso de diagnóstico de faltas indicaria que a amostra foi classificada corretamente como não tendo uma falta.

FP significa falso positivo, no caso de diagnóstico de faltas indicaria que a amostra foi classificada erroneamente para um determinado tipo de falta

FN significa falso negativo, no caso de diagnóstico de faltas indicaria que a amostra foi classificada não tendo um tipo de falta que na verdade teria.

A métrica acurácia é definida como:

$$Acurácia = \frac{TP + TN}{TP + TN + FP + FN} \quad (15)$$

A acurácia é uma das métricas mais utilizadas, já que divide o total de acertos pelo total de valores presentes na base de dados, e diz quanto o modelo acertou das classificações possíveis, então sua interpretação é a mais intuitiva.

A métrica precisão é definida por:

$$Precisão = \frac{TP}{TP + FP} \quad (16)$$

Esta métrica responde a seguinte pergunta: Qual a proporção de identificações positivas foi realmente correta? No caso de diagnóstico de faltas a métrica indica quantos casos previstos como determinada falha, realmente eram daquela falha.

A métrica sensibilidade (recall) é definida por:

$$Sensibilidade = \frac{TP}{TP + FN} \quad (17)$$

Esta métrica responde a seguinte pergunta: qual proporção de positivos foi identificada corretamente? Ou seja, com ela pode-se saber quão bom o modelo é em prever positivos (a classe que se quer prever), e no caso de diagnóstico de faltas, se uma amostra pertence a determinado tipo de falha.

A métrica F_score é definida por:

$$F_score = \frac{2TP}{2TP + FN + FP} \quad (18)$$

Esta métrica representa um balanço entre a precisão e a sensibilidade.

2.5 Estado da Arte de Redes Neurais Artificiais aplicadas ao problema de diagnóstico de faltas em Transformadores.

O problema de diagnóstico de falha em transformadores de potência tem sido alvo de vários artigos e trabalhos de pesquisadores, utilizando diferentes técnicas de inteligência computacional, sendo em sua maioria as RNA's e suas variantes.

Diversos pesquisadores utilizaram as RNAs para a classificação de falhas em transformadores, entre os quais destacam-se os trabalhos de (HAO; CAI-XIN, 2007) e (CASTRO; MIRANDA; LIMA, 2011). Entre as motivações para o uso das RNAs é citado sua capacidade de aproximação de funções bem como de adaptação a problemas de classificação.

No trabalho apresentado por (MIRANDA; CASTRO; LIMA, 2012) foi apresentado uma nova abordagem para o problema de classificação de faltas em transformadores, juntamente com o algoritmo “Mean Shift”. Um conjunto de redes neurais autoassociativas é treinado, objetivando-se que cada uma destas esteja “sintonizada” com um tipo de falha e então estas são dispostas em uma estrutura competitiva seguindo o princípio: a que conseguir maior correspondência com a amostra de entrada será a estrutura vencedora e indicará a condição do transformador. Os autores conseguiram 100% (cem por cento) de acurácia em seu modelo para base de teste, utilizando o banco de dados TC10 em conjunto com a técnica “Mean Shift” para a criação de mais 1400 dados virtuais que foram utilizados no processo de treinamento do sistema.

Na pesquisa apresentada em (WANG *et al.*, 2016) é mostrada uma metodologia para classificação de falhas em transformadores usando CSAE (*Continuous Sparse Auto Encoders*). CSAE pode ser considerado uma técnica de aprendizado de máquina chamada “manifold”, que se propõe a reduzir a dimensionalidade de dados não lineares e, para isso, se utiliza de uma

unidade Gaussiana Estocástica em sua função de ativação, ou seja, um termo a mais na equação que tem por benefícios modificar a direção do gradiente e prevenir o *overfitting*. Os autores utilizaram 2 camadas intermediárias de CSAE para fazer a extração de características e uma camada *backpropagation* na saída para performar a classificação. Os dados utilizados foram do banco TC10 e o resultado alcançado foi de 93,6% em média.

Em (XU *et al.*, 2019) é proposto um sistema de DBSPA (*Deep Belief Sparse Autoencoder*) para a classificação de faltas em transformadores, composto por camadas de DBSPA, responsáveis por executar a extração de características e uma camada de *backpropagation* na saída responsável por fazer a classificação. DBSPA são modelos de arquitetura consistentes de várias redes autoassociativas (*autoencoders*) e performam um aprendizado não supervisionado. No modelo proposto, o treinamento se dá na arquitetura inteira, com a atualização de todos os pesos de acordo com o erro gerado pela camada supervisionada de Backpropagation de saída. Os autores se utilizam do banco de dados TC10 bem como de dados provenientes de companhias elétricas chinesas e a rede proposta de DBSAE atingiu a acurácia média de 95,4% .

A abordagem em (MIRANDA; CASTRO, 2005) investiga a aplicação de um TFRENN (*Transparent Fuzzy rule extraction from neural network*) no diagnóstico de falhas em transformadores. Os autores “traduzem” a inferência de uma rede neural em um conjunto de expressões linguísticas familiares aos seres humanos por meio da lógica Fuzzy e com isso propõe uma nova tabela IEC melhorada. Utilizando os bancos de dados TC10, provenientes de literatura e da Companhia Elétrica do Pará (CELPA), os autores chegaram no resultado de 100% de acurácia para o sistema TFRENN (com 7 casos ambíguos) e 95,91% para a tabela IEC aprimorada.

Na pesquisa de (SOUAHLIA; BACHA; CHAARI, 2013) se propõe o uso de um modelo de classificação baseado em SVM (*Support Vector Machine*) para o diagnóstico de falhas. Os autores criaram uma arquitetura com 6 SVM's, cada um com a responsabilidade de diferenciar entre cada um dos tipos de falhas nos transformadores. Foi utilizado um banco de dados proveniente da companhia de eletricidade e gás da Tunísia e, para o treinamento, foi utilizado as razões do método de Rogers e de Doernenburg, chegando na conclusão que o primeiro apresenta uma maior acurácia, de 85%.

Para (LI *et al.*, 2016) o SaE-ELM (*Self adaptative evolutionay extreme learning machine*) combinado com “arctangent transform” (AT) é um modelo bem sucedido para a classificação de faltas em transformadores. O modelo combina os benefícios da teoria de “Evolução diferencial” (DE) com o modelo “aprendizado de máquina extremo” (ELM), sendo

os pesos do vetor de entrada do ELM otimizados usando a metodologia “algoritmo auto adaptativo de evolução diferencial” (SaE), derivada da teoria de DE, ao invés de randomicamente ou manualmente o fazer. Além disso, para lidar com dados polarizados, muito grandes ou muito pequenos, os autores empregaram o método AT para performar a normalização. Os dados utilizados provêm de duas fontes: Um compilado de literaturas anteriores e dados provenientes de uma companhia elétrica local. Com isso, os autores chegaram na acurácia de 79,17%.

Na pesquisa apresentada em (DAI *et al.*, 2017) é utilizado um modelo de DBN (*Deep Belief Network*) para o problema de diagnóstico de falhas em transformadores. DBN é um tipo de rede neural profunda composta por várias camadas de RBM (*Restricted Boltzmann Machine*) capaz de extrair as características relevantes de um conjunto de dados. A RBM é utilizada pela sua capacidade de representar grandes e múltiplas características dos dados de entrada e na camada de saída, há um classificador. Os dados são provenientes de publicações anteriores bem como da Companhia Nacional de Rede Elétrica da China. A acurácia atingida com o modelo foi de 89,2%.

Em (YANG *et al.*, 2021) é proposto um modelo de DSAE (*Double-stage autoencoder*) para o problema de classificação de faltas em transformadores. O DSAE é a junção de dois autoencoders (SAE – *Sparse Autoencoder*) com uma camada classificadora (Softmax) na saída e tem o objetivo de eliminar a influência de imperfeições nos dados no processo de diagnóstico, tais como: Dados não balanceados, dados não rotulados e erros de diagnóstico devido à idade de amostras. No modelo proposto, o DSAE objetiva diagnosticar os dados não balanceados, onde foi adicionado um critério de “tolerância ao envelhecimento” no SAE para reduzir o erro de diagnóstico por amostras velhas e o SDCN (*Advanced sparse deep clustering network*) é uma ferramenta de clusterização que objetiva diminuir o impacto de amostras não rotuladas. Os autores utilizam um banco de dados que congrega informações reais de coletas DGA (Dissolved Gas Analysis) e atingiu a acurácia de 98% em 0,2 segundos de análise.

Diversos algoritmos de RNAs têm sido apresentados na literatura como boas soluções para o problema, além de propor modificações para superar problemas, como o da falta de entendimento sobre o processo por de trás das RNAs em (CASTRO; MIRANDA, 2005) ou implementações para melhorar o processo de análise das condições de saúde do equipamento, como apresentado em (KIM *et al.*, 2020).

3 METODOLOGIA

3.1 Introdução

Como já mencionado, um transformador em serviço está sujeito a uma variedade de possíveis falhas térmicas e elétricas, sendo a detecção e eliminação destas de grande importância para o bom funcionamento desses equipamentos. Como apresentado no capítulo 2, um dos métodos mais utilizados para esse fim é o método de análise dos gases dissolvidos em óleo (DGA), que está em constante evolução e aperfeiçoamento de técnicas.

Seguindo a linha de aperfeiçoamento, várias técnicas de inteligência computacional, como as redes neurais artificiais, vêm sendo aplicadas ao problema de diagnóstico de faltas em transformadores. As RNAs reduzem o problema de diagnóstico de faltas a um processo de associação de variáveis de entrada (concentrações dos gases) e saída (tipo de falta).

Considerando as capacidades das RNAs, principalmente a de ser aproximadora universal de funções, este capítulo tem como objetivo apresentar a metodologia utilizada para o desenvolvimento do sistema proposto para diagnóstico de faltas incipientes em transformadores baseado na análise dos gases dissolvidos no óleo. Serão descritas as etapas realizadas para o desenvolvimento do sistema para classificação de faltas baseado em um sistema de redes Autoassociativas Empilhadas.

3.2 Base de dados de transformadores com faltas

Para o desenvolvimento do sistema proposto neste trabalho foram utilizados dados provenientes da base TC 10 de equipamentos faltosos, apresentada em (DUVAL; DEPABLA, 2001). Adicionalmente, uma base de dados proveniente de literaturas passadas empregada em (CASTRO, 2004) também foi utilizada. A base de dados TC 10 foi a base de dados utilizada para criação da tabela 6 apresentada no capítulo 2, sendo um dos métodos mais utilizados pelas empresas para diagnóstico de transformadores. O apêndice A apresenta todos os dados utilizados neste trabalho.

A base de dados contém um total de 296 amostras, sendo que cada amostra é formada pelas concentrações dos gases Hidrogênio (H_2), Metano (CH_4), Etano (C_2H_6), Etileno (C_2H_4) e Acetileno (C_2H_2), assim como pelo tipo de falta do transformador para estas concentrações. Segundo (MIRANDA; CASTRO; LIMA, 2012), é comum observar, em muitos estudos presentes na literatura que utilizam a base TC 10, que os casos T2 (Falha térmica com $300^\circ\text{C} <$

$T < 700^{\circ}\text{C}$) e T3 (Falha térmica com $T > 700^{\circ}\text{C}$), são agrupados em um único tipo de falta ($T > 300^{\circ}\text{C}$), e isto se deve ao fato do pequeno número de amostras para estas faltas na base de dados, o que seria inadequado para o processo de criação dos classificadores. Seguindo a mesma linha dos trabalhos apresentados na literatura, neste trabalho foram adotadas as faltas segundo nomenclatura apresentada na Tabela 8. Os dados foram divididos em 2 subconjuntos: de treinamento (71% do valor total de amostras) e o de teste (29% do valor total de amostras).

Tabela 8 – Tipos de Faltas do banco de dados

Tipo de Falta	Nomenclatura adotada	Número de amostras	
		Treinamento	Teste
Falha térmica $T < 300^{\circ}\text{C}$	T1	61	16
Falha térmica $T > 300^{\circ}\text{C}$	T2	52	16
Descargas parciais	PD	12	3
Descargas de baixa energia	DL	26	11
Descargas de alta energia	DH	60	39

3.3 Sistema de diagnóstico de faltas via rede Autoassociativas Empilhadas

3.3.1 Definição das entradas e saídas do sistema

Para o desenvolvimento do sistema classificador de faltas, diversas configurações de entradas (combinações das razões entre gases) foram testadas. O teste de diversas configurações se deve ao fato de um dos objetivos específicos deste trabalho que visa um estudo para verificação de outras possibilidades de combinação entre as razões de gases, além das que são comumente utilizadas pelos métodos convencionais, que possam trazer resultados de acurácia superior aos métodos convencionais. Para este estudo foram consideradas todas as razões apresentadas em (KIM *et al.*, 2013), o qual propôs 10 razões entre os gases chave que melhor diferenciavam os tipos de falha, de uma maneira gráfica. A Tabela 9 apresenta todas as razões de gases que foram testadas no nosso estudo.

Tabela 9 – Razões de gases

Código	Razão
R1	$\frac{CH_4}{H_2}$
R2	$\frac{C_2H_2}{C_2H_4}$
R3	$\frac{C_2H_2}{CH_4}$
R4	$\frac{C_2H_6}{C_2H_2}$
R5	$\frac{C_2H_4}{C_2H_6}$
R6	$\frac{C_2H_4}{CH_4}$
R7	$\frac{C_2H_6}{CH_4}$
R8	$\frac{C_2H_4}{H_2}$
R9	$\frac{C_2H_6}{H_2}$
R10	$\frac{C_2H_2}{H_2}$

Para o caso da saída do sistema classificador, os tipos de faltas foram codificados de tal forma que pudessem ser utilizados como saída da estrutura de redes autoassociativas empilhadas. A Tabela 10 apresenta a codificação utilizada para as faltas.

Tabela 10 – Código para saídas da rede neural

Tipo de Falta	Codificação de entrada	Codificação de saída
T1	1	[10000]
T2	2	[01000]
PD	3	[00100]
DL	4	[00010]
DH	5	[00001]

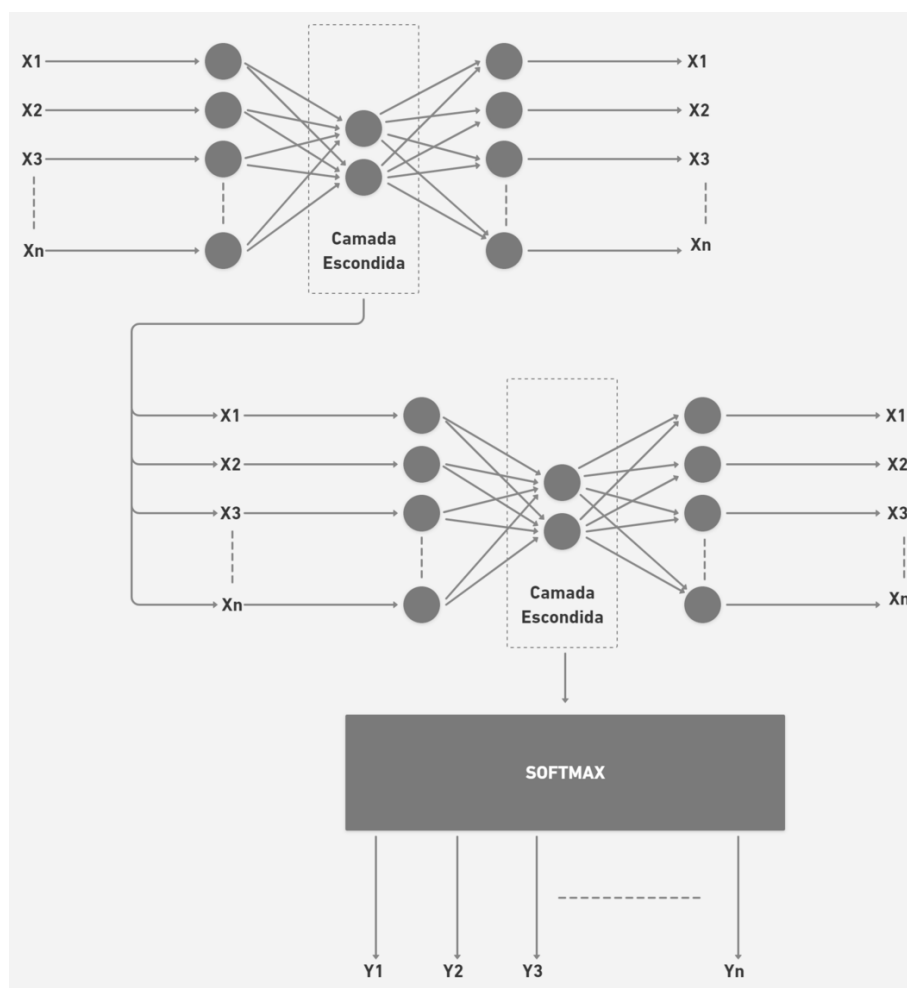
A base de dados de entrada passou por etapas de condicionamento e normalização. Baseado em um pré-conhecimento da tabela 6, todas as razões de gases com valores acima de

4, foram arredondadas para 4. Todas as divisões por zero também foram substituídas por 4, seguindo método apresentado pelos autores em (LI; WU; WU, 2011), e as indeterminações (zero dividido por zero) foram substituídas por zero.

3.3.2 Estrutura da rede Autoassociativa empilhada para o diagnóstico de faltas

Para desenvolvimento do sistema classificador diversas estruturas de redes autoassociativas empilhadas foram testadas. A Figura 13 apresenta a estrutura geral proposta para a classificação de faltas baseada nas redes Autoassociativas Empilhadas.

Figura 13 – Estrutura proposta



Fonte: Autor

Diversas topologias para estrutura de redes autoassociativas empilhadas também foram testadas, onde variou-se o número de redes autoassociativas empilhadas e para cada rede

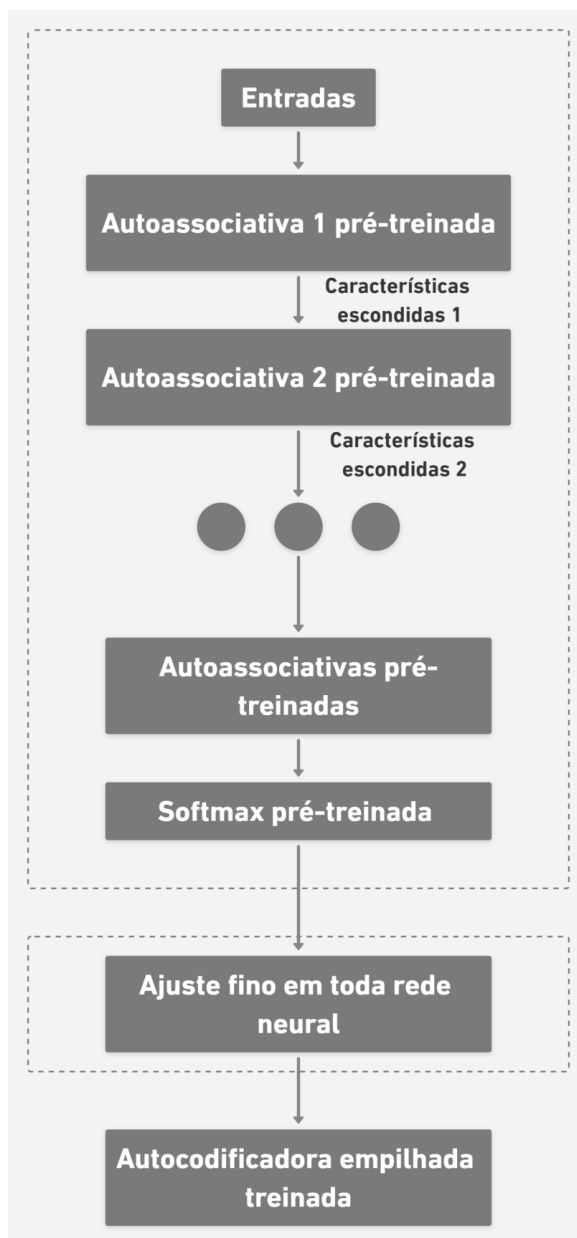
autoassociativa treinada, variou-se o número de neurônios na camada escondida codificadora. Todas as redes autoassociativas usaram a função sigmoideal como função de ativação dos neurônios.

O processo de treinamento da estrutura de redes autoassociativas empilhadas foi baseado em duas etapas, sendo que na primeira etapa as redes autoassociativas foram treinadas individualmente para que se chegasse a um vetor de características extraídas a partir da camada escondida codificadora de cada rede e na segunda etapa as autoassociativas foram empilhadas com a camada softmax para treinamento do sistema classificador.

De acordo com o esquema da Figura 14, a primeira rede autoassociativa recebe como entrada/saída as razões dos gases, sendo treinada com o objetivo de minimizar o erro médio quadrático entre saída e entrada da rede. Após treino da primeira rede, uma segunda rede autoassociativa é treinada, utilizando como entrada/saída os valores obtidos pela camada escondida codificadora da primeira autoassociativa já treinada. Foi definido inicialmente que a estrutura de redes autoassociativas empilhadas a serem testadas teriam no máximo 3 redes autoassociativas, sendo que após o treino das autoassociativas, inicia-se então o treino de toda a estrutura classificadora proposta como apresentada na Figura 13, onde as autoassociativas são empilhadas junto com a camada classificadora softmax. Nesta fase é realizado o treinamento da estrutura, buscando melhores resultados de classificação, considerando os dados da base de teste. A análise de desempenho foi baseada nas métricas de acurácia, f-score e sensibilidade e também através da análise da matriz de confusão.

Após o treino da softmax para classificação, toda a estrutura de rede autoassociativa empilhada passa ainda por um processo de ajuste fino dos parâmetros, buscando ainda melhorar o desempenho do sistema, considerando base de teste. A Figura 14 apresenta um esquemático de todo processo realizado para desenvolvimento do sistema classificador de faltas.

Figura 14 – Processo de treinamento de uma autoassociativa empilhada



Fonte: adaptado de (QI *et al.*, 2017)

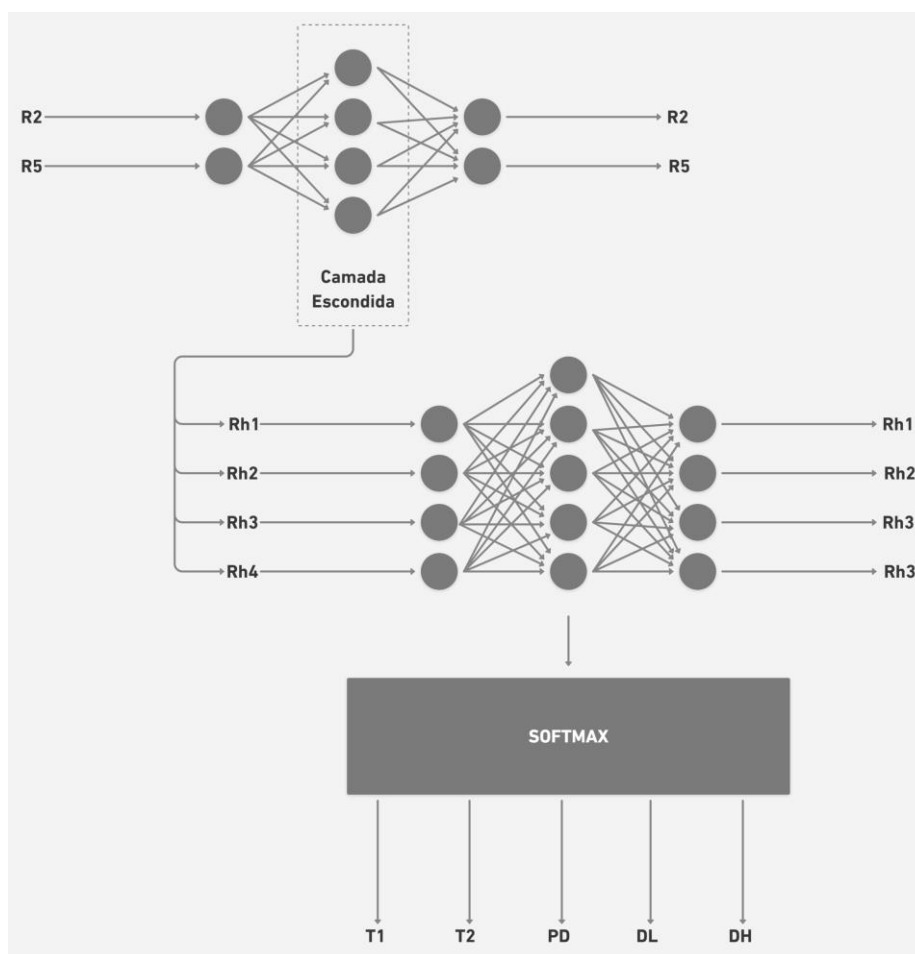
Para treinamento da estrutura de redes autoassociativas empilhadas foi utilizado o toolbox de redes neurais profundas do Matlab. Para treino das redes autoassociativas e camada softmax foi utilizado o algoritmo de aprendizado *Scaled Conjugate Gradient (SCG)* (JYOTI; RAWAT, 2017). Todos os treinos das autoassociativas utilizaram como critério de parada o número de épocas, que foi definido para todas como 600 épocas. O apêndice B apresenta o código MATLAB utilizado para a criação da rede neural. Para todas as estruturas testadas, diversas inicializações dos pesos foram realizadas, buscando evitar o problema de mínimo local.

4 RESULTADOS

4.1 Sistema de diagnóstico baseado em redes Autoassociativas empilhadas

Após todo o processo de treino das diversas estruturas de redes autoassociativas empilhadas, como descrito na metodologia, o melhor resultado encontrado, considerando desempenho na base de teste, foi para estrutura com 2 entradas (razões R2 e R5), 2 redes autoassociativas, com 4 neurônios na camada intermediária codificadora da primeira autoassociativa e 5 neurônios na camada intermediária codificadora da segunda autoassociativa. A Figura 15 apresenta a estrutura final do sistema classificador.

Figura 15 – Estrutura do modelo proposto



Fonte: Autor

A acurácia obtida, considerando a base de teste, foi de 100%, ou seja, do total de padrões de teste, não houve erros na classificação das faltas, como pode ser verificado através

da matriz de confusão da Figura 16. O sistema apresentou 100% de sensibilidade e f-score. Nenhum falso negativo ou falso positivo.

Figura 16 – Matriz de confusão

Classes Preditas	16 18.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	16 18.8%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	0 0.0%	3 3.5%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	0 0.0%	0 0.0%	11 12.9%	0 0.0%	100% 0.0%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	39 45.9%	100% 0.0%
	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%
Classes Reais						

Fonte: Autor

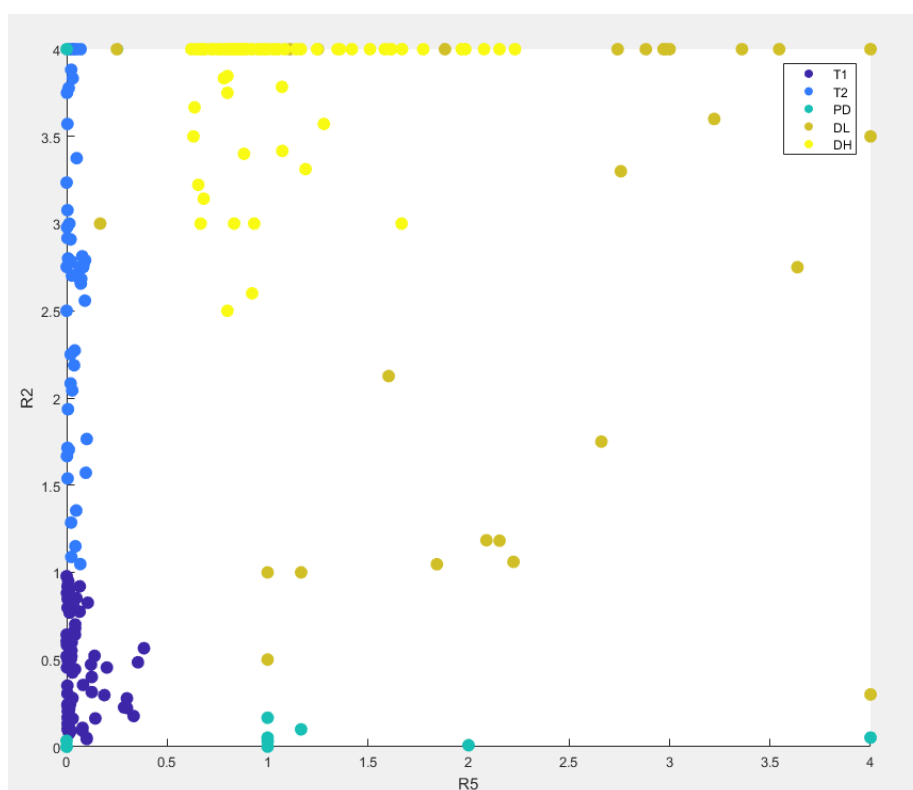
As duas razões de gases que foram utilizadas como entrada do sistema, são duas das razões utilizadas, entre as 3, no método tradicional IEC 60599. Os resultados mostram que somente com estas duas razões, já se consegue dividir bem o espaço entre as classes com o uso da estrutura proposta.

As duas razões de gases que foram utilizadas como entrada do sistema, são duas das razões utilizadas, entre as 3, no método tradicional IEC 60599. Os resultados mostram que somente com estas duas razões (normalizadas), já se consegue dividir bem o espaço entre as classes com o uso da estrutura proposta.

Em (KIM *et al.*, 2013), os autores indicam que as razões R2 e R5 (sem normalização) podem ser utilizadas para fazer uma boa separação entre as falhas térmicas na classe de $300^{\circ}\text{C} < T < 700^{\circ}\text{C}$, classe esta que não é utilizada neste trabalho e a classe de $T > 700^{\circ}\text{C}$, bem como para fazer a distinção entre PD, DL e DH. Todavia, como evidenciado no mesmo estudo, esta separação não é imune a falhas e algumas classificações podem ser ambíguas, havendo a

necessidade de complementação com outras razões. Para o nosso trabalho, esta dificuldade foi superada utilizando somente estas duas razões. Considerando a base de treino e teste usada no nosso trabalho, a Figura 17 demonstra a organização espacial das amostras de acordo com os valores de R2 e R5 (com normalização adotada), onde é possível perceber que estas duas razões permitem uma boa separação entre todas as 5 classes (existem poucos casos que se sobrepõem), sendo que a estrutura de redes autoassociativas empilhadas proposta teve a capacidade de aprender uma função não-linear que conseguiu promover a acurácia de acerto do sistema de 100%.

Figura 17 – Gráfico em função das razões



Fonte: Autor

4.2 Estado comparativo com método IEC 60599 e Sistemas baseados em Rede Neural Perceptron de Múltiplas camadas

Para fins de comparação foram usados 3 métodos de diagnósticos de faltas em transformadores, são eles: Tabela IEC 60599, rede 1, que consiste em uma MLP treinada considerando como entrada as 3 razões usadas na tabela IEC 60599 e uma rede 2, que consiste em uma MLP treinada considerando como entrada as razões R2 e R5. As duas MLPs foram treinadas a partir da mesma base de dados utilizada para desenvolvimento do sistema proposto.

A rede 1 foi treinada para comparação do sistema proposto com uma estrutura já consolidada na literatura para diagnóstico de faltas. A rede 2 foi implementada para comparação com a MLP usando as mesmas entradas do sistema proposto e teve como objetivo mostrar a eficiência da estrutura de rede proposta para o problema de diagnóstico. Para o treinamento, as entradas para as redes 1 e 2 também foram normalizadas de acordo com o mesmo critério adotado para a autoassociativas da estrutura empilhada. Diversos treinos foram realizados para rede 1 e 2 para obter o melhor desempenho considerando a base de teste. A Tabela 11 apresenta os resultados obtidos para comparação com o sistema proposto.

Tabela 11 – Comparação do modelo proposto com outros métodos

Método	Taxa de acerto na base de teste
Modelo proposto	100%
Tabela IEC	80%
Rede 1	97%
Rede 2	97%

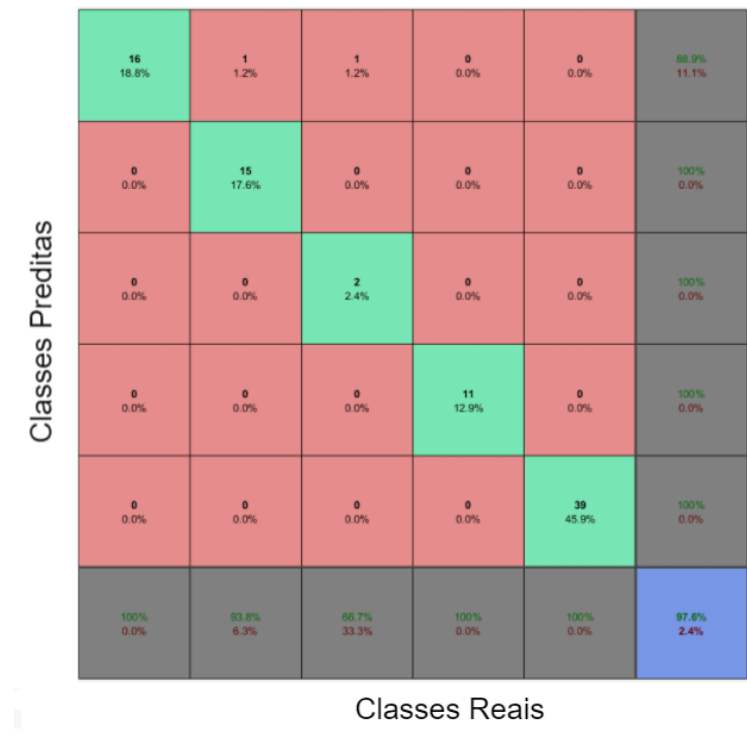
A Figura 18 apresenta a matriz de confusão para a rede 1 e a Figura 19 a matriz de confusão para rede 2.

Figura 18– Matriz de confusão da rede 1



Fonte: Autor

Figura 19 – Matriz de confusão da rede 2



Fonte: Autor

Dos resultados obtidos temos que o sistema proposto foi o que obteve melhor resultado de classificação na base de teste. O sistema conseguiu com apenas duas razões como entrada para estrutura, separar as classes de tal forma que problemas de casos Não-identificados (não-decisão) ou classificação errada encontrados no uso da Tabela IEC 60599 foram eliminados. No caso dos resultados obtidos para Tabela IEC, a mesma obteve, para base de teste, 7 casos de erro de classificação e 10 casos de falta não identificada.

Para os casos da rede 1 e rede 2, observa-se que as duas têm os mesmos resultados de acurácia, tendo obtido 2 casos de erro de classificação. As duas redes apresentam resultados superior a tabela IEC 60599, se mostrando também adequadas para o problema de diagnóstico de transformadores, entretanto, quando comparadas com o modelo proposto, este se mostrou mais eficiente para separação das classes através da função não-linear aprendida através da base de treinamento.

5 CONCLUSÃO

Este trabalho de conclusão de curso apresentou os resultados obtidos na pesquisa de aplicação de redes neurais autoassociativas empilhadas para o problema de diagnóstico de faltas em transformadores, utilizando a análise dos gases dissolvidos em óleo (DGA). Para o desenvolvimento do sistema foi utilizada a base de dados TC 10 de equipamentos faltosos usada na publicação da norma IEC 60599, bem como dados provenientes de literatura.

O bom resultado alcançado no modelo proposto, de 100% na taxa de acerto na base de teste, mostra que o sistema apresenta uma boa generalização, ou seja, responde adequadamente para dados não utilizados na fase de treinamento. Além disso, é importante salientar que dentro dos resultados têm-se casos que a rede neural conseguiu fornecer diagnóstico correto enquanto o método IEC 60599 não reconheceu o tipo de falha, problema de “não-decisão” discutido anteriormente.

Sabe-se que este assunto é de grande relevância para o setor elétrico e grandes indústrias que possuem parques industriais, visto que ainda se utilizam muitos transformadores a óleo em suas instalações e as falhas são constantes e, por muitas vezes, não mapeadas, devido à dificuldade em extrair as informações úteis sobre o estado operativo dos equipamentos. Nesse cenário, a utilização de redes neurais surge como um aliado no processo de manutenção preditiva para o diagnóstico rápido e eficiente destas faltas.

Como trabalhos futuros, pretende-se acrescentar no modelo de diagnóstico o estado de “sem falha”, um desafio devido a pouca quantidade de dados disponíveis atualmente. Além disso, pode-se utilizar uma metodologia de criação de dados virtuais para mitigar tal problema, bem como melhorar a precisão deste e outros modelos de classificação de faltas em transformadores.

REFERÊNCIAS BIBLIOGRÁFICAS

APTE, Sandip; SOMALWAR, Rahul; WAJIRABADKAR, Ashwini. “Incipient Fault Diagnosis of Transformer by DGA Using Fuzzy Logic”. *In: 2018 IEEE INTERNATIONAL CONFERENCE ON POWER ELECTRONICS, DRIVES AND ENERGY SYSTEMS (PEDES)*, 2018. **2018 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)**. [S. l.: s. n.], 2018. p. 1–5.

BENMAHAMED, Y.; TEGUAR, M.; BOUBAKEUR, A. Diagnosis of Power Transformer Oil Using PSO-SVM and KNN Classifiers. *In: 2018 INTERNATIONAL CONFERENCE ON ELECTRICAL SCIENCES AND TECHNOLOGIES IN MAGHREB (CISTEM)*, 2018. **2018 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM)**. [S. l.: s. n.], 2018. p. 1–4.

CASTRO, Adriana Rosa Garcez. Knowledge extraction from artificial neural networks : Application to transformer incipient fault diagnosis. [s. l.], 2004. Disponível em: <https://repositorio-aberto.up.pt/handle/10216/12369>. Acesso em: 12 ago. 2021.

CASTRO, A.R.G.; MIRANDA, V. Knowledge discovery in neural networks with application to transformer failure diagnosis. **IEEE Transactions on Power Systems**, [s. l.], v. 20, n. 2, p. 717–724, 2005.

CASTRO, Adriana R. Garcez; MIRANDA, Vladimiro; LIMA, Shigeaki. Transformer fault diagnosis based on autoassociative neural networks. *In: 2011 16TH INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEM APPLICATIONS TO POWER SYSTEMS*, 2011. **2011 16th International Conference on Intelligent System Applications to Power Systems**. [S. l.: s. n.], 2011. p. 1–5.

DAI, Jiejie *et al.* Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network. **IEEE Transactions on Dielectrics and Electrical Insulation**, [s. l.], v. 24, n. 5, p. 2828–2835, 2017.

DHINI, Arian *et al.* Intelligent Fault Diagnosis for Power Transformer Based on DGA Data Using Support Vector Machine (SVM). *In: 2018 3RD INTERNATIONAL CONFERENCE ON SYSTEM RELIABILITY AND SAFETY (ICSRS)*, 2018. **2018 3rd International Conference on System Reliability and Safety (ICSRS)**. [S. l.: s. n.], 2018. p. 294–298.

DUVAL, M.; DEPABLA, A. Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases. **IEEE Electrical Insulation Magazine**, [s. l.], v. 17, n. 2, p. 31–41, 2001.

FAIZ, Jawad; SOLEIMANI, Milad. Dissolved gas analysis evaluation in electric power transformers using conventional methods a review. **IEEE Transactions on Dielectrics and Electrical Insulation**, [s. l.], v. 24, n. 2, p. 1239–1248, 2017.

GUO, Chenxi; DONG, Ming; WU, Zhanyu. Fault Diagnosis of Power Transformers Based on Comprehensive Machine Learning of Dissolved Gas Analysis. *In: 2019 IEEE 20TH INTERNATIONAL CONFERENCE ON DIELECTRIC LIQUIDS (ICDL)*, 2019. **2019 IEEE 20th International Conference on Dielectric Liquids (ICDL)**. [S. l.: s. n.], 2019. p. 1–4.

HAO, Xiong; CAI-XIN, Sun. Artificial Immune Network Classification Algorithm for Fault Diagnosis of Power Transformer. **IEEE Transactions on Power Delivery**, [s. l.], v. 22, n. 2, p. 930–935, 2007.

IEC 60599:1999, MINERAL OIL-IMPREGNATED ELECTRICAL EQUIPMENT IN SERVICE - GUIDE TO THE INTERPRETATION OF DISSOLVED AND FREE GASES ANALYSIS. . [S. l.]: International Electrotechnical Commission, 1999. International Standard. Disponível em: <https://webstore.iec.ch/publication/2582#additionalinfo>. .

IEEE GUIDE FOR THE INTERPRETATION OF GASES GENERATED IN OIL-IMMERSED TRANSFORMERS. **IEEE Std C57.104-2008 (Revision of IEEE Std C57.104-1991)**, [s. l.], p. 1–36, 2009.

IRUNGU, G. K.; AKUMU, A. O.; MUNDA, J. L. Comparison of IEC 60599 gas ratios and an integrated fuzzy-evidential reasoning approach in fault identification using dissolved gas analysis. *In*: 2016 51ST INTERNATIONAL UNIVERSITIES POWER ENGINEERING CONFERENCE (UPEC), 2016. **2016 51st International Universities Power Engineering Conference (UPEC)**. [S. l.: s. n.], 2016. p. 1–6.

JYOTI; RAWAT, Anil Kumar. Scale conjugate gradient based learning applied to handwritten digit classification. *In*: 2017 CONFERENCE ON INFORMATION AND COMMUNICATION TECHNOLOGY (CICT), 2017. **2017 Conference on Information and Communication Technology (CICT)**. [S. l.: s. n.], 2017. p. 1–6.

KIM, Sunuwe *et al.* A Semi-Supervised Autoencoder With an Auxiliary Task (SAAT) for Power Transformer Fault Diagnosis Using Dissolved Gas Analysis. **IEEE Access**, [s. l.], v. 8, p. 178295–178310, 2020.

KIM, Sung-wook *et al.* New methods of DGA diagnosis using IEC TC 10 and related databases Part 1: application of gas-ratio combinations. **IEEE Transactions on Dielectrics and Electrical Insulation**, [s. l.], v. 20, n. 2, p. 685–690, 2013.

LI, Shuaibing *et al.* Interpretation of DGA for transformer fault diagnosis with complementary SaE-ELM and arctangent transform. **IEEE Transactions on Dielectrics and Electrical Insulation**, [s. l.], v. 23, n. 1, p. 586–595, 2016.

LI, Xiaohui; WU, Huaren; WU, Danning. DGA Interpretation Scheme Derived From Case Study. **IEEE Transactions on Power Delivery**, [s. l.], v. 26, n. 2, p. 1292–1293, 2011.

MIRANDA, V.; CASTRO, A.R.G. Improving the IEC table for transformer failure diagnosis with knowledge extraction from neural networks. **IEEE Transactions on Power Delivery**, [s. l.], v. 20, n. 4, p. 2509–2516, 2005.

MIRANDA, Vladimiro; CASTRO, Adriana R. Garcez; LIMA, Shigeaki. Diagnosing Faults in Power Transformers With Autoassociative Neural Networks and Mean Shift. **IEEE Transactions on Power Delivery**, [s. l.], v. 27, n. 3, p. 1350–1357, 2012.

MORAIS, Lorena dos Reis. Estrutura de redes neurais auto-associativas aplicadas ao processo de identificação de equipamentos elétricos em sistemas de monitoramento não intrusivo de cargas. **1 CD-ROM**, [s. l.], 2019. Disponível em: <http://repositorio.ufpa.br/jspui/handle/2011/12166>. Acesso em: 26 ago. 2021.

MORAIS, Lorena Dos Reis. **Sistema Inteligente Para Monitoramento e Diagnóstico De Transformadores Baseado na Análise Físico-química Do Óleo E Análise Dos Gases Dissolvidos Em Óleo**. 2008. Dissertação (Mestrado Acadêmico em Engenharia Elétrica) - Universidade Federal do Pará, Pará, 2008.

N'CHO, Janvier Sylvestre *et al.* Review of Physicochemical-Based Diagnostic Techniques for Assessing Insulation Condition in Aged Transformers. **Energies**, [s. l.], v. 9, n. 5, p. 367, 2016.

PALKE, Rahul; KORDE, Pragati. Dissolved Gas Analysis (DGA) to Diagnose the Internal Faults of Power Transformer by using Fuzzy Logic Method. *In: 2020 INTERNATIONAL CONFERENCE ON COMMUNICATION AND SIGNAL PROCESSING (ICCSP)*, 2020. **2020 International Conference on Communication and Signal Processing (ICCSP)**. [S. l.: s. n.], 2020. p. 1050–1053.

QI, Yumei *et al.* Stacked Sparse Autoencoder-Based Deep Network for Fault Diagnosis of Rotating Machinery. **IEEE Access**, [s. l.], v. 5, p. 15066–15079, 2017.

SAHRI, Zahriah; YUSOF, Rubiyah. Fault diagnosis of power transformer using optimally selected DGA features and SVM. *In: 2015 10TH ASIAN CONTROL CONFERENCE (ASCC)*, 2015. **2015 10th Asian Control Conference (ASCC)**. [S. l.: s. n.], 2015. p. 1–5.

SILVA, Ivan Nunes da. **Redes Neurais Artificiais para engenharia e ciências aplicadas: Fundamentos teóricos e aspectos práticos**. 2. ed. São Paulo: Artliber Editora Ltda, 2016.

SOUAHLIA, Seifeddine; BACHA, Khmais; CHAARI, Abdelkader. SVM-based decision for power transformers fault diagnosis using Rogers and Doernenburg ratios DGA. *In: 10TH INTERNATIONAL MULTI-CONFERENCES ON SYSTEMS, SIGNALS DEVICES 2013 (SSD13)*, 2013. **10th International Multi-Conferences on Systems, Signals Devices 2013 (SSD13)**. [S. l.: s. n.], 2013. p. 1–6.

WANG, Lukun *et al.* Transformer fault diagnosis using continuous sparse autoencoder. **SpringerPlus**, [s. l.], v. 5, p. 448, 2016.

XU, Zhong *et al.* Transformer Fault Diagnosis based on Deep Brief Sparse Autoencoder. *In: 2019 CHINESE CONTROL CONFERENCE (CCC)*, 2019. **2019 Chinese Control Conference (CCC)**. [S. l.: s. n.], 2019. p. 7432–7435.

YANG, Dongsheng *et al.* A Novel Double-Stacked Autoencoder for Power Transformers DGA Signals with Imbalanced Data Structure. **IEEE Transactions on Industrial Electronics**, [s. l.], p. 1–1, 2021.

APÊNDICE A

Este apêndice apresenta o banco de dados de equipamentos faltosos usado para treinamento e teste do sistema de diagnóstico apresentado neste trabalho.

A tabela A.1 apresenta os dados de teste e a tabela A.2 os dados de treinamento usados no sistema proposto. O banco de dados TC 10 é composto por amostras de equipamentos faltosos inspecionados em serviço usados na publicação IEC60599 e apresentados em (DUVAL; DEPABLA, 2001), além de um banco de dados proveniente de literaturas passadas usado em (CASTRO, 2004).

Tabela A.1 – Banco de dados usado no treinamento

Amostra	H2	CH4	C2H2	C2H4	C2H6	Tipo de Falta	Fonte
	305	538	0	101	157	T1	TC 10
	220	1660	0	1140	1880		
	93	194	0	27	52		
	12	17.6	0	4.5	4.6		
	43	50	0.5	12	27	T1	Literatura
	16.2	17	0.1	6.1	42		
	60	73	0.3	19	22		
	16	31	0.3	10	62		
	58	77	0.1	18	88		
	131	198	0.3	52	105		
	17	28	0.4	23	94		
	13	19	0.5	6.4	69		
	26	28	0.9	11	31		
	22	31	1.5	36	56		
	27	71	2.8	58	68		
	32	65	0.3	20	26		
	37	55	0.3	17	21		
	15	16	0.1	0.8	2		
	12	17	0.5	1.3	2.3		
	46	71	0.2	19	94		
	5	42	0.1	4.2	8.1		
	5	38	0.2	4.7	6.7		
	7	9.8	0.2	1.9	2.3		
	10.2	139	0.4	19	116		
	29.9	129	0.1	72	123		
	9	17	0.2	3.1	4		

11	13	0.2	1.6	5.1		
10	46	0.6	9.2	10		
20	100	0.1	20	150		
51	83	0.1	24	99		
20	37	0.1	4.2	7.6		
50	51	0.2	16	73		
20	49	0.1	6.3	9.9		
20	72	0.1	23	25		
20	42	0.1	4.6	9.1		
20	22	0.1	26	85		
110	150	0.1	50	110		
40	79	0.1	20	120		
24	62	0.3	63	79		
20	135	0.2	16	210		
20	86	0.1	26	43		
36	80	0.4	78	92		
43	65	0.1	67	76		
20	21	0.1	53	91		
20	37	0.1	4.7	8.9		
20	61	0.2	17	19		
20	33	0.1	3.9	6.5		
36	52	0.8	6.6	14		
19	32	0.5	3.6	6.9		
33	84	0.5	39	43		
7.8	140	0.1	16	135		
0.1	11	0.1	0.3	1.7		
10.7	31	0.6	7.6	70		
9.9	38	0.2	4.7	6.9		
5.5	36	0.2	5.8	7.2		
10	33	1.1	3.1	6.4		
10.5	32	0.1	11	65		
21.3	110	0.1	20	200		
29.8	30	0.1	23	97		
11.6	140	0.1	15	140		
12	25	0.1	7	15		
<hr/>					T2	Literatura
22	51	0.5	46	27		
42	120	4.5	330	110		
1700	3580	135	2700	800		

5	7.4	0.1	2.5	1.1		
23	68	1.3	47	23		
25	65	0.4	60	31		
45	51	1.8	120	43		
45	51	1.8	120	43		
7.1	8.4	0.3	3	1.7		
20	23	0.7	23	6		
10	11	0.7	9	3.2		
10	14	0.6	8.5	3.2		
10	17	0.2	4.2	3.1		
11	12	0.7	11	4.1		
11	12	0.9	11	4		
25	53	1.3	35	16		
39	47	0.5	120	39		
39	47	0.5	120	39		
10	14	1.1	13	4.7		
10	13	1.1	12	4.3		
10	13	1	11	4.3		
10	13	0.7	11	4		
20	46	0.1	24	14		
35	42	2.1	22	14		
20	22	0.2	10	4.8		
11	12	0.8	11	4.1		
100	230	5.4	270	120		
20	24	0.1	60	16		
20	47	0.1	20	13		
10	13	0.1	25	7		
44	74	2.5	128	44		
3.6	6.7	0.1	2.3	2		
5	11	0.1	4.5	3.5		
33.9	36	1.6	78	28		
33.9	36	1.6	78	28		
33.6	42	1.5	66	17		
<hr/>						
200	700	0	500	200	T2	TC 10
30	200	8	308	114		
465	3100	1	3360	1221		
495	1775	2	2480	326		
6709	10500	750	17700	1400		
290	966	57	1810	299		
101	184	10	243	32		

50	100	9	305	51		
3650	6730	191	9630	1570		
1040	2100	10	2720	579		
107	143	2	222	35		
78	259	0	640	117		
80	6129	0	2438	276		
507	1053	17	1440	297		
723	1988	32	3259	595		
416	695	0	867	74		
32930	2397	0	0	157	PD	Literatura
39	1.7	0.1	0.1	0.6		
37800	1740	8	8	249	PD	TC 10
92600	10200	0	0	0		
9340	995	7	6	60		
350	0	144	574	0		
350	0	144	574	0		
6600	1000	19	2	38		
88	9	0	0.03	0		
2240	168	0	0	25		
33046	619	0	2	58		
40280	1069	1	1	1060		
4	1	52	7	2	DL	TC 10
1900	285	7730	957	31		
35	6	482	26	3		
650	81	270	51	170		
210	22	7	6	6		
385	60	159	53	8		
4230	690	1180	196	5		
84	9	40	11	4		
1790	580	619	336	321		
57	24	30	27	2		
1000	500	500	400	1		
2240	157	45	45	90		
73	8	2	12	4		
5000	4000	2000	8000	2000		
2240	560	940	450	380		
117	16	58	18	5		
60	5	29	6	1		
890	110	700	84	3		
41	112	4536	254	0		

16000	4000	16000	8500	500		
60	10	4	4	4		
6870	1028	5500	900	79		
10092	5399	37565	6500	530		
24.3	15.7	29.8	11.2	6.4		
2240	360	828	169	25		
4480	560	896	403	380		
<hr/>						
1570	1110	1830	1780	175	DH	TC 10
260	215	277	334	35		
13500	6110	4040	4510	212		
34	21	56	49	4		
420	250	800	530	41		
310	230	760	610	54		
10000	6730	10400	7330	345		
1500	395	323	395	28		
20000	13000	57000	29000	1850		
620	325	244	181	38		
62	16	63	65	1.8		
13.8	1.4	0.8	1.2	0.1		
48.3	33	282	169	21		
20	5.6	51	61	2.1		
110	25	110	150	14		
32	11	2.3	3.2	0.6		
25	5.4	8.2	8.6	1		
20	8.3	26	32	3.2		
150	120	340	390	30		
<hr/>						
20	5	26	34	0.7	DH	Literatura
23	13	15	14	3.7		
182	164	520	633	64		
87	17	94	120	9.7		
28	7.8	2.5	3.1	0.7		
20	12	40	38	1.2		
20	4.3	1.2	1.5	0.4		
20	5.6	34	41	0.8		
20	3.6	81	81	0.9		
20	7.7	1.5	1.7	0.5		
20	5.4	50	66	2.6		
20	9.2	24	26	2.3		
20	3.6	1	1.2	0.4		
20	4.4	2.1	2.5	0.6		

20	11	55	31	7.1
56	14	35	52	1.6
26	7	10	15	2.5
32	6.7	18	29	4.6
20	2.4	0.2	0.3	0.1
20	2.6	26	36	0.3
20	3.5	42	48	0.7
20	4.5	18	27	0.6
20	3.1	8.4	13	0.3
20	5.7	53	80	2.7
20	3.3	22	25	0.6
20	7.9	1.5	2.2	0.7
20	5.2	1.8	2.3	0.6
73	25	85	100	4.8
72	8.6	5.3	8.4	2.4
20	7.7	3.7	4.5	1.1
20	8	4	5	1.3
17	9.7	6.4	5	1.4
26	5.4	1.9	2.9	0.9
25	3.6	2.4	2.9	0.5
22	2.9	1.6	2.1	0.5
29	17	29	13	1.9
16	10	6.3	5.3	1.6
30	19	27	13	2
12	11	4.4	4.1	1.2
18	5.7	1.2	1.3	0.5
1000	680	650	620	55

Tabela A.2 – Banco de dados usado no teste

Amostra	H2	CH4	C2H2	C2H4	C2H6	Tipo de Falta	Fonte
	6.2	6.6	0.2	0.7	3.1	T1	TC 10
	10.4	98	0.4	19	205		
	21.7	41	0.7	24	86		
	0.8	26	0.3	1	4.5		
	1	25	0.1	0.7	4.3		
	6.9	17	1	5	11		
	4.9	31	0.3	38	40		
	0.5	27	0.3	1.6	5.4		
	0.5	22	0.3	1	3.6		
	2.4	26	0.1	3.4	4		
	5	8.6	0.2	11	13		
	10.8	23	0.1	27	77		
	13	14	0.1	8.2	35		
	16	19	2	69	162		
	50	200	1	10	225		
	10	170	1	10	200		
	20	21	0.2	8.6	7.9	T2	Literatura
	34	39	0.7	98	35		
	20	22	0.3	4.4	4.2		
	49	68	0.6	140	48		
	20	57	0.1	55	33		
	48	91	3.6	340	90		
	98	160	0.3	140	47		
	20	48	0.2	29	17		
	231	3997	0	5584	1726		
	100	230	5.4	270	120		
	107	143	2	222	34		
	860	1670	40	2050	30		
	100	200	11	670	110		
	53	83	10	144	31		
	200	680	88	1600	190		
	1050	2400	11	1800	370		
	36036	4704	10	5	554	PD	TC 10
	8266	1061	0	0	22		
	1950	123	2	2	38		
	120	25	40	8	1	DL	TC 10

6454	2313	6432	2159	121		
78	20	28	13	11		
305	100	541	161	33		
1230	163	692	233	27		
645	86	317	110	13		
95	10	39	11	0		
595	80	244	89	9		
1330	10	182	66	20		
543	120	1880	411	41		
2177	1049	705	440	207		
<hr/>						
9474	4066	12997	6552	353	DH	Literatura
441	207	261	224	43		
64	24	190	120	19		
22	3	5	3	1		
64	24	190	120	19		
138	65	103	112	9		
22	3	5	3	1		
5.3	2.8	1.6	1.6	0.3		
6.2	1.4	1	1.2	0.1		
3.3	1.2	0.4	0.5	0.1		
1.2	1	0.5	0.5	0.1		
7.2	4.3	106	78	1		
18.4	7.7	97	97	2.2		
5	3	11	16	1		
150	63	220	220	15		
57	19	28	35	14		
19.1	2.8	33	38	0.6		
24	4.4	13	15	0.9		
14.8	9.1	56	26	3.7		
6	2.8	24	36	0.1		
14.1	4.4	18	25	0.1		
3.5	2.9	1.4	1.5	0.5		
11	4.3	0.7	1.1	0.3		
7.1	2.6	2.7	2.7	0.5		
6	2.6	28	33	0.6		
6	1.3	1.5	1.8	0.2		
31	20	103	95	10		
13	7.6	20	30	2.2		
13	2.1	0.4	0.5	0.2		
14	7.8	42	26	2.8		

37	7	22	28	5
6.1	5.6	3.8	2.4	0.5
29	15	68	100	2.3
15	1.7	0.5	0.6	0.1
19	11	290	180	4.7
74	25	120	140	14
8	1	0.4	0.5	0.1
22	7	67	68	2.5
150	67	250	280	22

APÊNDICE B

Este apêndice apresenta o código fonte do programa em MATLAB que foi usado para o desenvolvimento do sistema de diagnóstico apresentado neste trabalho.

```
%% Criando o xtreino
valores = []; % Aqui entram os valores do banco de dados de treino

% Agora vamos precisar fazer as razões e depois trata-las. A ordem é
a %seguinte:
% H2 - CH4 - C2H2 - C2H4 - C2H6

%Como queremos as razões R2 e R5 de (KIM et al., 2020), temos que
R2 = C2H2/C2H4,
R5 = C2H4/C2H6. Assim

R2 = valores(:,3) ./ valores(:,4); %C2H2/C2H4
R5 = valores(:,4) ./ valores(:,5); %C2H4/C2H6

%Agora, precisamos criar o vetor de treino, concatenando as razões
xtreino = horzcat (R2,R5);

%E agora precisamos tratar os valores

%Performamos o seguinte: Para valores > 4, substitui por 4
% Para divisão por zero, substitui por 4
% Para 0/0, substitui por 0

%Lidando com valores NaN - Ocorre quando temos 0/0
xtreino(isnan(xtreino)) = 0;

for i=1:numel(xtreino)
    if (xtreino(i) == Inf)%Lidando com valores Inf - Ocorre quando temos
x/0, um numero qualquer por zero
        xtreino(i) = 4;
    elseif (xtreino(i) > 4) %Lidando com valores acima de 4
        xtreino(i) = 4;
    end
end

%Arrumamos o xtreino
xtreino = xtreino';

%% Criamos os valores de saída para o treino
xsaida = []; % Aqui entram os valores das respectivas classes dos dados de
treinamento
xsaida = xsaida';

%% Agora vamos repetir os processos para o vetor de testes, Criando xteste

valores_teste = []; % Aqui entram os valores do banco de dados de teste
```

```

% Agora vamos precisar fazer as razões e depois trata-las. A ordem é a
seguinte:
% H2 - CH4 - C2H2 - C2H4 - C2H6

%Como queremos as razões R2 e R5 de DuvalNew, temos que R2 = C2H2/C2H4,
R5 = C2H4/C2H6. Assim

R2_teste = valores_teste(:,3) ./ valores_teste(:,4); %C2H2/C2H4
R5_teste = valores_teste(:,4) ./ valores_teste(:,5); %C2H4/C2H6

%Agora, precisamos criar o vetor de treino, concatenando as razões
xteste = horzcat (R2_teste,R5_teste);

%E agora precisamos tratar os valores

%Performamos o seguinte: Para valores > 4, substitui por 4
% Para divisão por zero, substitui por 4
% Para 0/0, substitui por 0

%Lidando com valores NaN - Ocorre quando temos 0/0
xteste(isnan(xteste)) = 0;

for i=1:numel(xteste)
    if (xteste(i) == Inf)%Lidando com valores Inf - Ocorre quando temos x/0,
um numero qualquer por zero
        xteste(i) = 4;
    elseif (xteste(i) > 4) %Lidando com valores acima de 4
        xteste(i) = 4;
    end
end

%Arrumando xteste
xteste = xteste';

%% Criando desejado
desejado = []; % Aqui entram os valores das respectivas classes dos dados
de teste

desejado = desejado';

%% Criando os vetores no padrão Softmax

xtreino_softmax = []; % Aqui entram os valores das respectivas classes dos
dados de treinamento, no formato de 0's e 1's, ex: [00001]
xtreino_softmax = xtreino_softmax';

xteste_softmax = []; % Aqui entram os valores das respectivas classes dos
dados de teste, no formato de 0's e 1's, ex: [00001]
xteste_softmax = xteste_softmax';

%% Criando a Stacked Autoencoder

```

```

autoenc1 = trainAutoencoder(xtreino,4, ...
    'MaxEpochs',600, ...
    'L2WeightRegularization',0.004, ...
    'SparsityRegularization',4, ...
    'SparsityProportion',0.15, ...
    'ScaleData', false);
feat1 = encode(autoenc1,xtreino);

autoenc2 = trainAutoencoder(feat1,5, ...
    'MaxEpochs',600, ...
    'L2WeightRegularization',0.004, ...
    'SparsityRegularization',4, ...
    'SparsityProportion',0.15, ...
    'ScaleData', false);
feat2 = encode(autoenc2,feat1);

softnet = trainSoftmaxLayer(feat2,xtreino_softmax,'MaxEpochs',400);
view(softnet)

stackednet = stack(autoenc1,autoenc2,softnet);
view(stackednet)

%RESULTADO SEM FINE TUNING
%y = stackednet(xteste);
%plotconfusion(xteste_softmax,y);

%RESULTADO COM FINE TUNING
stackednet = train(stackednet,xtreino,xtreino_softmax);
y = stackednet(xteste);
plotconfusion(xteste_softmax,y);

```