

Universidade Federal de Campina Grande
Centro de Engenharia Elétrica e Informática
Unidade Acadêmica de Engenharia Elétrica
Grupo de Sistemas Elétricos

Redes Neurais Aplicadas no Diagnóstico de Descargas Parciais

Trabalho de Conclusão de Curso

Aluno: Francisco Alex de Oliveira Gouveia

Orientador: Edson Guedes da Costa

Campina Grande, Outubro 2008



Biblioteca Setorial do CDSA. Fevereiro de 2021.

Sumé - PB

Título:

**Redes Neurais Aplicadas no Diagnóstico de Descargas
Parciais**

Aluno:

Francisco Alex de Oliveira Gouveia

Francisco Alex de Oliveira Gouveia

Orientador:

Edson Guedes da Costa

Campina Grande, Outubro 2008.

Agradecimentos

Agradeço a minha família pela oportunidade em especial a meu pai pelo grande incentivo. A meus amigos Álvaro, Glauber, Rafael, Philipe, Luciano, Antonio, Luciana, Helder e Jobson, pela grande contribuição e amizades ao longo dos anos de graduação. Ao professor Edson Guedes e Max Gleison Gonçalves Neri pela oportunidade de trabalho e aos funcionários do Laboratório de Alta Tensão Eduardo, Francisco e Adriano. A todos aqueles que no momento não foram lembrados, mas que tiveram grande importância nessa caminhada rumo obtenção do título de engenheiro electricista.

SUMÁRIO

Resumo.....	6
1 Revisão Bibliográfica.....	7
1.1 Descargas Parciais.....	7
1.2 Operadores Estatísticos.....	11
1.2.1 Assimetria.....	13
1.2.2 Curtose.....	14
1.2.3 Centro de Distribuição.....	16
1.2.4 Correlação Cruzada.....	16
1.2.5 Desvio Padrão.....	17
1.2.6 Fator de Descarga.....	17
1.3 Redes Neurais Artificiais.....	18
1.3.1 Modelo do Neurônio.....	19
1.3.2 Tipos de Função de Ativação.....	20
1.3.3 Arquitetura de Redes Neurais.....	22
1.3.4 Processo de Aprendizagem.....	25
1.3.5 Redes Perceptron de Múltiplas Camadas.....	25
1.3.6 Treinamento da Rede.....	26
1.3.7 Algoritmo de Retropropagação.....	27
1.3.8 Taxa de Aprendizagem.....	31
1.3.9 Critério de Parada.....	31
1.3.10 Algoritmo Resilient Propagation.....	32
2 Materiais e Métodos.....	34
3 Simulações e Resultados.....	38
3.1 Primeira etapa.....	38
3.2 Segunda etapa.....	40
3.3 Terceira etapa.....	43
3.4 Quarta etapa.....	44
4 Bibliografia.....	47

Lista de Tabelas

Tabela 1 - Código binário de cada tipo de DP.....	37
Tabela 2 - Resultado dos testes da terceira etapa.....	44
Tabela 3 - Resultado dos testes da quarta etapa.....	45

Lista de Figuras

Figura 1 - Modelo da cavidade.....	8
Figura 2 - Comportamento da tensão e corrente na descarga.....	8
Figura 3 - Descarga parcial utilizada na análise de fase.....	10
Figura 4 - Sinais de descarga parcial utilizados na análise de tempo (NERI, 2008).....	11
Figura 5 - Modelo de distribuição normal.....	12
Figura 6 – Exemplo de um histograma.....	12
Figura 7 - Distribuição com assimetria igual à zero.....	13
Figura 8 - Distribuição com assimetria positiva.....	14
Figura 9 - Distribuição com assimetria negativa.....	14
Figura 10 - Curtose igual à zero.....	15
Figura 11 - Curtose positivo.....	15
Figura 12 - Curtose negativo.....	15
Figura 13 - Sinal com mesmo valor de assimetria e Curtose.....	16
Figura 14 - Sinais com correlação cruzada diferentes.....	17
Figura 15 - Modelo de um neurônio artificial.....	19
Figura 16 - Função de limiar brusco.....	21
Figura 17 - Função limiar lógico.....	21
Figura 18 - Função de sigmóide.....	22
Figura 19 - Arquitetura de uma rede de camada única.....	23
Figura 20 - Arquitetura de uma rede de múltiplas camadas.....	24
Figura 21 - Arquitetura de uma rede recorrente de camada única.....	24
Figura 22 - Arranjo do sistema de medição de descargas parciais com capacitor de acoplamento e bobina de bloqueio (NERI, 2008).....	35
Figura 23 - Semi-ciclo do sinal de DP dividido em janelas.....	36
Figura 24 - Parte do arquivo da base de dados.....	37
Figura 25 - Estrutura de RNA utilizada no treinamento.....	39
Figura 26 - Desempenho do erro primeira etapa sem validação.....	39
Figura 27 - Desempenho do erro primeira etapa com validação.....	40
Figura 28 - Desempenho do erro segunda etapa sem validação (máximos).....	41
Figura 29 - Desempenho do erro segunda etapa com validação (máximos).....	42
Figura 30 - Desempenho do erro segunda etapa sem validação (médias).....	43
Figura 31 - Desempenho do erro segunda etapa com validação (médias).....	43

Resumo

O isolamento elétrico constitui uma parte importante de um sistema elétrico, fundamental para sua segurança e funcionamento eficaz. Devido sua importância existe a necessidade de realizar periodicamente diagnósticos que possam caracterizar o estado do isolamento dos equipamentos, identificando possíveis anomalias. Dentre as técnicas de avaliação existente, a medição de descargas parciais mostrou-se ao longo de vários anos de estudo uma ferramenta confiável na detecção de anomalias. Os estudos a cerca das descargas parciais visam aperfeiçoar os sistemas de aquisição, desenvolver metodologias de ensaio e técnicas de análise dos sinais obtidos. Devido às características, o sinal de descarga parcial apresenta formato específico para cada tipo e nível de anomalia avaliada, o que representa uma espécie de impressão digital do defeito e pode ser utilizado para distinguir padrões com relação ao estado do isolamento. A análise dos sinais de descargas é uma tarefa subjetiva que depende necessariamente da experiência do responsável que ira avaliar as características do sinal e compara-lá com algum exemplo da literatura. O fato motiva a utilização de ferramentas matemáticas com objetivo de extrair características que possam representar o formato do sinal de descarga, para assim facilitar a classificação de padrões das descargas. Por se tratar de uma tarefa com caráter subjetivo, a classificação pode fazer uso de técnicas de processamento de informações baseadas no funcionamento do cérebro humano, chamadas redes neurais artificiais. As redes servem para realizar a identificação e classificação de padrões do isolamento estudado. A classificação e a identificação dos padrões são conseguidas mediante processo de aprendizagem onde diversos padrões de descargas parciais são apresentados à rede neural e, a partir de algoritmos de aprendizagem, a rede é dotada da capacidade de generalizar o fenômeno estudado. O trabalho consistiu em utilizar sinais de descargas parciais obtidos em ensaios de laboratório, dos quais extraíram-se características por meio de operadores estatísticos que serviram para caracterizar o comportamento da descarga. Os dados fomaram uma base de conhecimento para treinamento de algumas estruturas de rede neural. Alguns testes foram realizados com as características extraídas dos sinais visando quantificar o grau de importância de cada operador estatístico no desempenho da rede.

1 Revisão Bibliográfica

1.1 Descargas Parciais

O estudo de descargas parciais (DPs) é uma ferramenta poderosa para avaliação do nível de degradação do isolamento de equipamentos de alta tensão. Desde década de 1920 vem sendo utilizado no diagnóstico de equipamentos, mas somente na década de 1950 ganhou força motivada pelo uso intensivo de isolamentos poliméricos, que são susceptíveis à degradação provocada pela ocorrência de descargas parciais, logo, existe a necessidade de monitorar o processo de fabricação, com a finalidade de minimizar a existência de impurezas e micro-cavidades (bolhas internas no isolamento), que favorecem o aparecimento de DPs.

Descargas parciais são descargas elétricas pequenas que curto-circuitam parcialmente a isolação. Estas descargas podem aparecer em cavidades gasosas, provenientes de falhas na fabricação, em meio isolante sólido ou líquido, ou ao longo de uma superfície ao redor de pontas. As DPs são decorrentes da ionização de uma determinada região do dielétrico. A ionização é provocada pelo campo elétrico, cuja intensidade depende da forma do eletrodo, das características do dielétrico e da tensão aplicada. Devido à não-uniformidade do campo elétrico, pequenas regiões do isolamento podem ser submetidas a solicitações superiores à sua rigidez dielétrica, resultando em descargas parciais.

As DPs são divididas em dois tipos de acordo com o local de origem:

Descarga corona – Ocorre em um meio gasoso, externo ao material isolante. As descargas corona acontecem principalmente em torno de condutores ou superfícies pontiagudas.

Descarga interna – São provocadas geralmente pela existência de bolhas preenchidas com gás. Tem origem no processo de fabricação ou por estresse no material ao longo da vida.

Os requisitos necessários para existência de descargas parciais são (SILVA, 2005):

1. É necessária a presença de um gás.

2. É necessária a presença de um campo elétrico que exceda o limite de isolamento suportável pelo gás.
3. É necessária a presença de pelo menos 1 elétron livre para que o processo seja desencadeado.

Um modelo simplificado de um dielétrico com uma cavidade preenchida com ar Figura 1 (a) pode ser utilizado para compreensão dos parâmetros utilizados na avaliação das DPs. No caso de uma tensão ser aplicada aos terminais A e B do modelo o arranjo é representado pelo circuito da Figura 1 (b) (COSTA e NEVES, 1996).

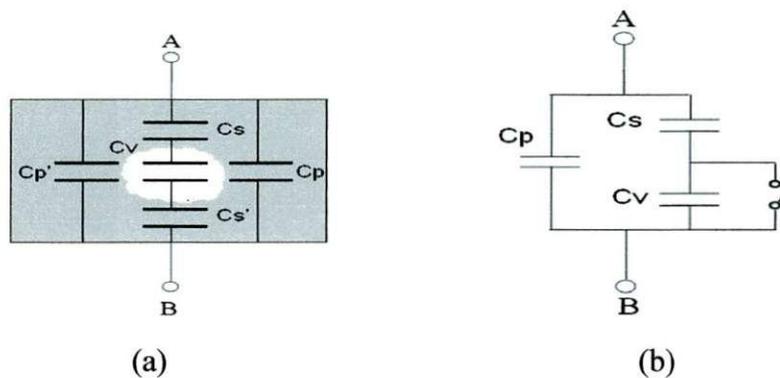


Figura 1 - Modelo da cavidade.

A capacitância da cavidade é representada por C_v , a capacitância do material que está em série com a cavidade é representada por C_s e C_p é a capacitância do resto do material. A chave utilizada representa o curto-circuito no momento da descarga. Para melhor compreender o processo de descarga é apresentado na Figura 2 o comportamento da tensão e corrente na cavidade da descarga.

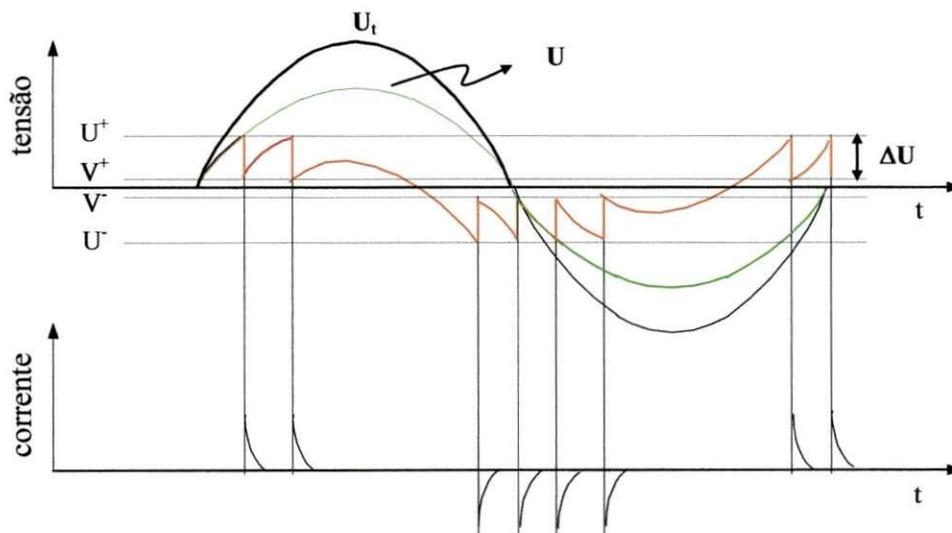


Figura 2 - Comportamento da tensão e corrente na descarga.

Onde, $U_i(t)$ é a tensão de alimentação alternada no material sob ensaio; $U(t)$ a tensão aparente que deveria existir na cavidade, no caso da descarga elétrica não acontecer; U^+ e U^- são as tensões disruptivas da cavidade para as duas situações de polaridade da tensão de alimentação; V^+ e V^- são as tensões residuais existentes nas cavidades, após a ocorrência das descargas elétricas, devido à transferência de cargas. Geralmente, tais descargas aparecem como pulsos elétricos com duração menor que 1 μ s (NERI, 2008).

A ocorrência de DPs é acompanhada por eventos que podem ser utilizados no estudo e caracterização da integridade do isolamento analisado. Abaixo são listados os tipos eventos (MOTA, 2002):

1. Geração de luz visível e radiação ultravioleta.
2. Produção de gás ozônio e oxigênio.
3. Ruído audível e não audível.
4. Formação de componentes químicos na parede dos dielétricos.
5. Geração de calor localizado e perda de energia elétrica no sistema.
6. Erosão mecânica da superfície devido ao bombardeamento iônico.
7. Rádio interferência.

A medição direta de DPs é uma tarefa difícil devido ao grande volume dos equipamentos com relação às dimensões dos defeitos que originam as descargas. Entretanto existem métodos de medição indireta com objetivo de quantificar as cargas envolvidas durante a descarga.

No momento da descarga o capacitor C_s , na Figura 1, sofre uma diminuição ΔU na tensão. Para restaurar o nível de tensão, uma carga ΔQ é drenada do capacitor C_p , que está em paralelo a C_v , e dos capacitores presentes no circuito externo que alimentam o sistema. A carga é encontrada pela expressão (1) (MOTA, 2002).

$$\Delta Q = \Delta U \frac{C_v \cdot C_s}{C_v + C_s} \quad (1)$$

O valor de ΔQ é diretamente proporcional à variação de tensão na cavidade. A variação ΔU é uma medida da magnitude ou intensidade da descarga, diretamente relacionada à intensidade da degradação causada pela DP. A afirmação fica clara quando

observado a equação de energia dissipada em cada pulso de descarga (MOTA, 2002), onde U representa a tensão disruptiva (2).

$$\Delta W \approx C_v \cdot \Delta U \left(U - \frac{1}{2} \Delta U \right). \quad (2)$$

A descarga parcial no interior da cavidade ocasiona a circulação de uma corrente no circuito externo do dielétrico. Devido ao fato de conseguir detectar somente parte da corrente transferida na descarga, a medição da corrente durante uma DP é uma medição indireta da carga transferida durante a descarga parcial, que por sua vez é uma medição indireta de ΔU . Sendo ΔU proporcional à energia dispendida pela descarga e transferida às paredes do dielétrico, logo está diretamente relacionado ao nível de degradação provocada pela descarga.

A medição das DPs em laboratório visa captar os sinais de corrente durante a descarga através de equipamentos detectores ou osciloscópios, para que sejam realizadas análises com respeito as suas características.

O sinal de DP tem um comportamento diferente para cada tipo e intensidade da anomalia provocada pelas descargas parciais, tornando o sinal de DP uma espécie de impressão digital do defeito.

Atualmente dois métodos de análise de descargas parciais encontram grande aplicação, método de análise de fase e o método de análise no tempo. A análise de fase tem a finalidade de caracterizar a distribuição de ocorrências dos pulsos de DPs com relação ao ângulo da fase da senóide da tensão de alimentação. De acordo com a posição, nível e geometria das DPs é possível identificar os sinais. Na Figura 3 é apresentado um sinal de descarga parcial utilizado na análise fase.

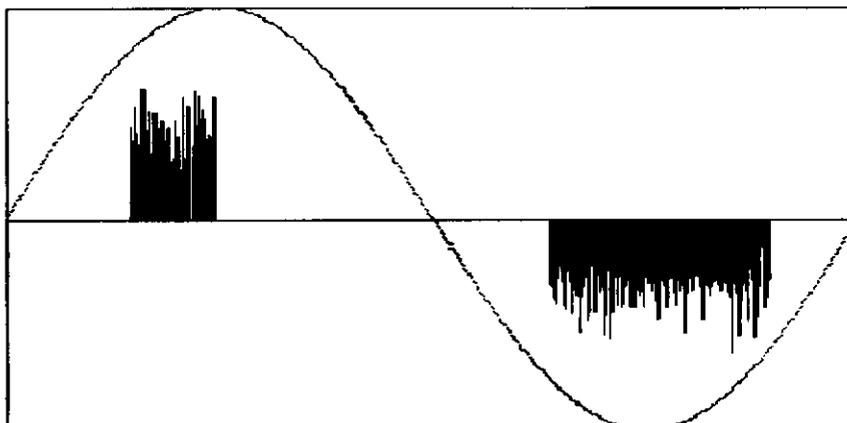


Figura 3 - Descarga parcial utilizada na análise de fase.

A análise de tempo tem a finalidade de avaliar parâmetros relacionados à forma da onda como tempo de subida, tempo de meia onda, e busca relaciona as grandezas com o estado do dielétrico avaliado. O método requer um sinal com menor distorção possível, logo se necessita de sistema de aquisição cuja faixa de passagem atenda os requisitos. Na Figura 4 é apresentada à tela de um osciloscópio com sinais de DP utilizado na análise de tempo.

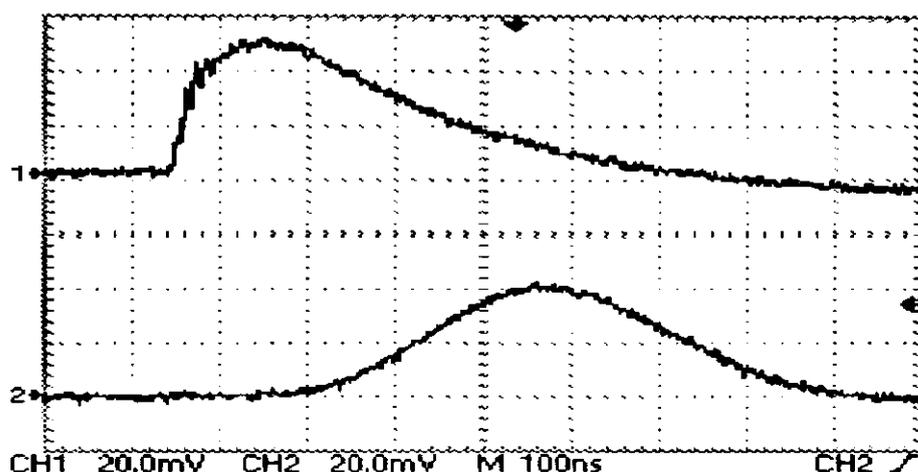


Figura 4 - Sinais de descarga parcial utilizados na análise de tempo (NERI, 2008).

1.2 Operadores Estatísticos

O formato da descarga é alterado por fatores como intensidade do campo elétrico, tipo de cavidade e clima o que dificulta a caracterização dos sinais de DPs. A análise de descargas parciais consiste em avaliar o seu formato e associá-lo a algum caso existente na literatura. Uma alternativa utilizada na análise consiste em realizar o processamento do sinal para extração de informações que caracterizem o formato das DPs, diminuindo assim as dificuldades de diagnóstico.

As DPs são fenômenos de físicos de natureza aleatória que depende de muitas variáveis. As características dos sinais viabiliza a aplicação de modelo estatístico de distribuição de probabilidade no estudo de DPs. A idéia consiste em quantizar a similaridade da distribuição dos sinais de DPs com relação ao modelo de probabilidade normal padrão, traduzindo as variações de forma em valores numéricos. A função de distribuição de probabilidade normal tem o formato sino, mostrado na Figura 5, onde o ponto de máximo encontra-se na média e a largura depende do desvio padrão do conjunto de dados.

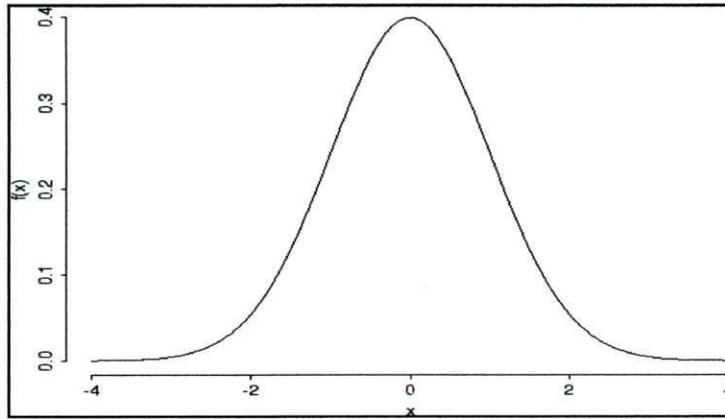


Figura 5 - Modelo de distribuição normal.

A distribuição das DPs pode ser comparada com a distribuição normal quando se realiza o agrupamento das descargas de um semi-ciclo em um histograma, que representa frequência dos dados em intervalos, o formato do gráfico tem semelhança com uma distribuição normal como mostrado na Figura 6. À medida que os intervalos do histograma são diminuídos seu formato aproxima-se da curva em vermelho que indica a distribuição normal.

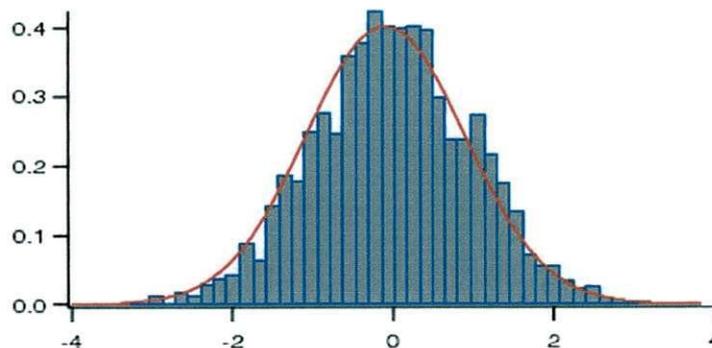


Figura 6 – Exemplo de um histograma.

No caso do estudo das DPs alguns trabalhos utilizaram operadores estatísticos para extrair informações sobre o formato da descarga para o reconhecimento de padrões, a exemplo de (KREUGER, 1991). Esta técnica consiste na medida e análise das distribuições estatísticas das ocorrências de DPs em função da amplitude e do ângulo de fase de ocorrência das mesmas (SILVA, 2005).

Na quantização do sinal DPs os operadores estatísticos que avaliavam o formato da distribuição das descargas com relação ao modelo distribuição normal, o resultado encontrado indicava o grau de semelhança entre as distribuições. Os operadores estatísticos utilizados e suas características serão apresentados a seguir.

1.2.1 Assimetria

A assimetria é o deslocamento uma distribuição em relação a sua média. No caso da distribuição ser simétrica o seu valor é zero. Para o caso da distribuição ser assimétrica à esquerda o operador tem o valor positivo. Para o caso da assimetria ser à direita o valor do operador é negativo.

No caso de descargas parciais o operador assimetria é encontrado da seguinte forma,

$$m_3 = \sum \frac{(X - \bar{X})^3}{n}; \quad (3)$$

$$m_2 = \sum \frac{(X - \bar{X})^2}{n}; \quad (4)$$

$$sk = \frac{m_3}{m_2 \sqrt{m_2}}, \quad (5)$$

X representa a amplitude da cada pulso de descarga, \bar{X} é a média aritmética da distribuição e n o número de descargas ocorrida.

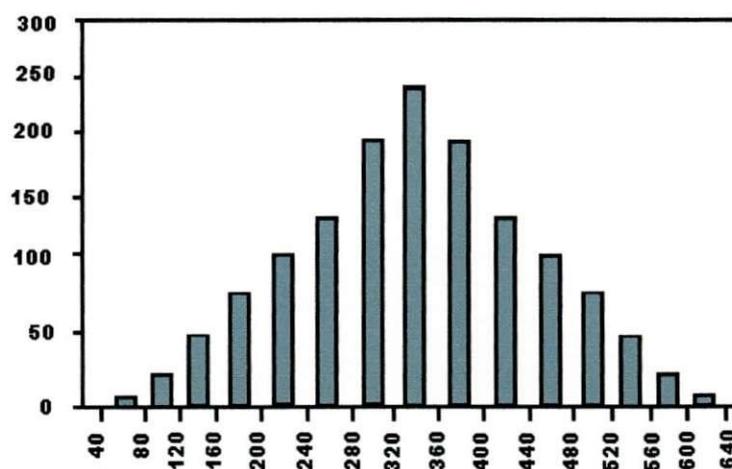


Figura 7 - Distribuição com assimetria igual à zero.

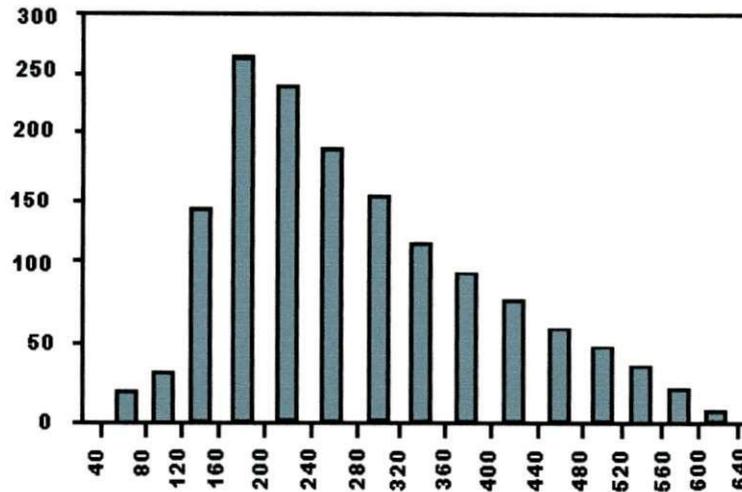


Figura 8 - Distribuição com assimetria positiva.

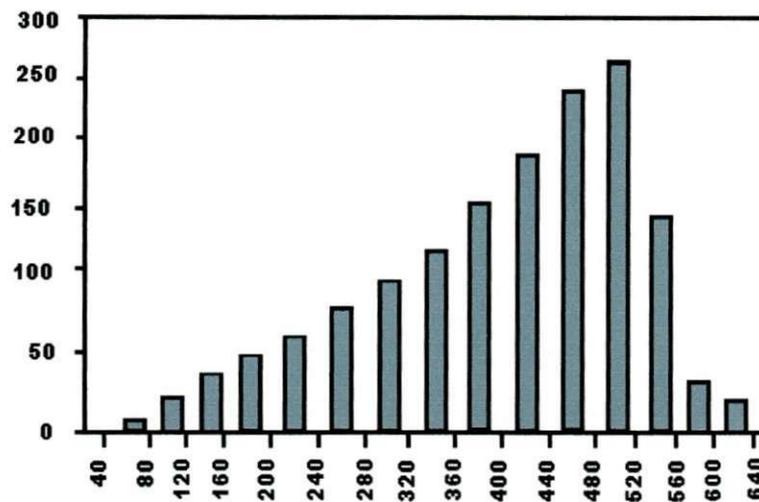


Figura 9 - Distribuição com assimetria negativa.

1.2.2 Curtose

Constitui um operador que avalia o espalhamento da distribuição com relação à média, tomando como referência a distribuição normal. No caso da distribuição ser igual a normal padrão o valor do operador é zero. Para o caso da distribuição ser estreita ficando muito próxima da média o valor do operador é positivo e no caso de ser mais espalhado o operador tem valor negativo. O cálculo da curtose é efetuado por:

$$m_4 = \sum \frac{(X - \bar{X})^4}{n}; \quad (6)$$

$$ku = \left(\frac{m_4}{m_2^2} \right) - 3. \quad (7)$$

Nas figuras abaixo se verificam os valores do Curtose para algumas distribuições.

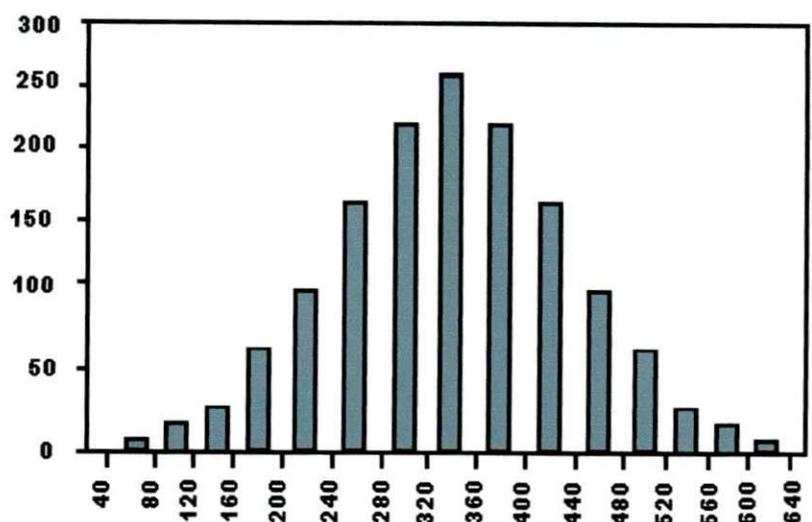


Figura 10 - Curtose igual à zero.

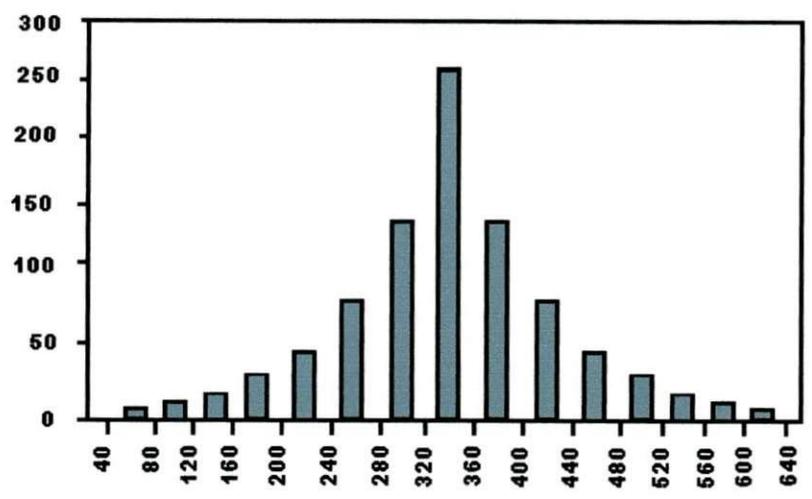


Figura 11 - Curtose positivo.

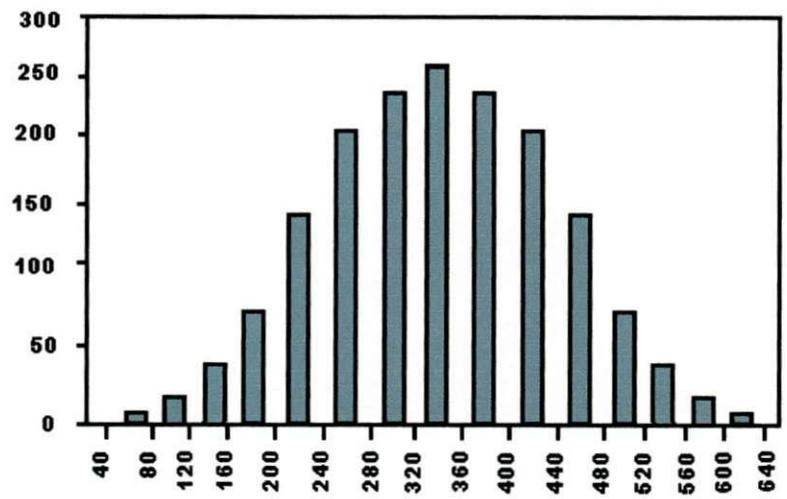


Figura 12 - Curtose negativo.

1.2.3 Centro de Distribuição

O operador tem o papel de determinar o instante do período da senóide em que a média da distribuição está localizada. Ele tem grande importância, pois existem distribuições diferentes com mesmo valor de assimetria e Curtose e com valores de centro de distribuição diferentes com no exemplo da Figura 13.

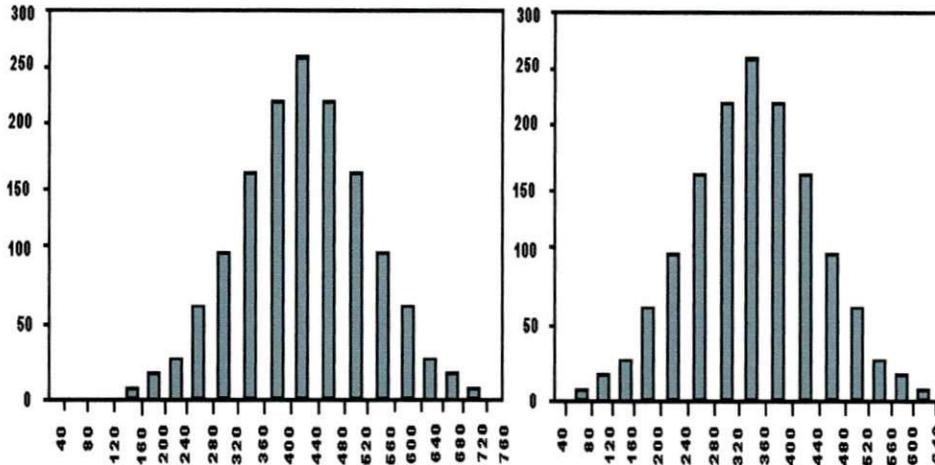


Figura 13 - Sinal com mesmo valor de assimetria e Curtose.

O centro de distribuição é encontrado pela seguinte expressão:

$$\bar{t} = \frac{1}{n} \sum f_j t_j . \quad (8)$$

O valor de f_j representa amplitude dos pulsos de descarga, t_j corresponde ao instante de tempo de ocorrência dos pulsos. O resultado encontrado é o instante, no período da senóide, em que está localizada a média aritmética da distribuição. O operador é utilizado nos dois semi-ciclo de uma senóide.

1.2.4 Correlação Cruzada

A correlação cruzada é um operador que avalia a diferença entre o formato da descarga no semi-ciclo positivo e negativo. No caso das descargas nos semi-ciclos tiverem o mesmo formato o valor do operador será igual a 1, caso sejam completamente diferentes o valor é zero. Na Figura 14, observa-se o exemplo de sinais o primeiro sinal

tem correlação cruzada próximo de zero devido à grande diferença entre as descargas do semi-ciclos, já o segundo sinal devido à similaridade tem valor próximo de 1.

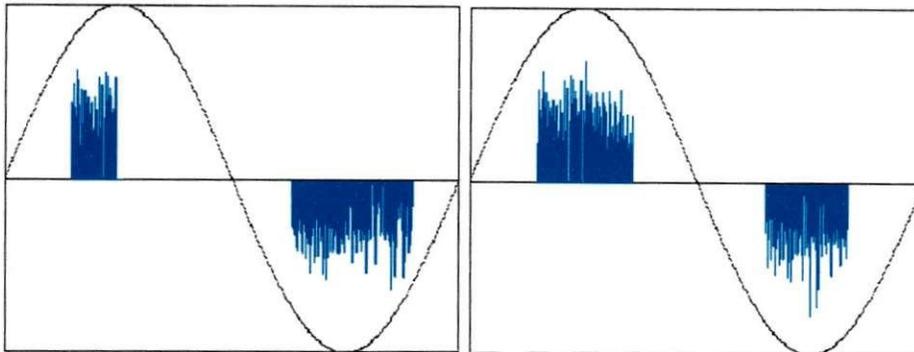


Figura 14 - Sinais com correlação cruzada diferentes.

A correlação cruzada é calculada da seguinte forma:

$$cc = \frac{\sum x_i y_i - \sum x_i \sum y_i}{n} \cdot \frac{1}{\sqrt{\left[\frac{\sum x_i^2 - (\sum x_i)^2}{n} \right] \left[\frac{\sum y_i^2 - (\sum y_i)^2}{n} \right]}} \quad (9)$$

As variáveis x_i e y_i representam a amplitude das descargas nos semi-ciclos positivo e negativo.

1.2.5 Desvio Padrão

O desvio padrão fornece uma estimativa do espalhamento da distribuição com relação à média. Ele é calculado da seguinte forma:

$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{n-1}} \quad (10)$$

1.2.6 Fator de Descarga

Fator de descarga indica a relação entre os níveis médios de descarga no semi-ciclo positivo e negativo. As variáveis Q é a soma das descargas nos semi-ciclo e N é o número de descarga no semi-ciclo, Q é dado por:

$$Q = \frac{\left(\frac{Q_s^-}{N^-}\right)}{\left(\frac{Q_s^+}{N^+}\right)}. \quad (11)$$

1.3 Redes Neurais Artificiais

O cérebro humano funciona como um computador de alta complexidade com a capacidade de organizar seus constituintes, os neurônios, de modo a realizar tarefas complexas e muito mais rápidas do que um computador digital. Com intuito de desenvolver um sistema que conseguisse contemplar as características do processamento do cérebro humano, estudos foram iniciados visando desenvolver um sistema que tivesse como base as estrutura neurais do cérebro humano, conhecidas como redes neurais artificiais.

As redes neurais artificiais (RNAs) são modelos de processamento de informação inspirados na estrutura física natural do cérebro humano. Elas são sistemas inerentemente paralelos, não lineares e distribuídos, compostos por unidades de processamento simples (neurônios) que têm características particulares. Estas unidades estão dispostas em estruturas interligadas por um grande número de conexões (sinapses), onde as conexões são associadas a pesos, que servem como fator de ponderação das entradas recebidas por cada unidade de rede. Redes neurais artificiais são geralmente implementadas utilizando componentes eletrônicos ou são simuladas em softwares através de processamento digital.

Uma das principais características das redes neurais é a adaptação ou aprendizagem através do método chamado de algoritmo de aprendizagem. O algoritmo utiliza exemplos na simulação da rede e modificação dos pesos sinápticos para alcançar o seu objetivo.

As RNAs encontram grande aplicação em diversos ramos da ciência principalmente em atividades relacionadas com reconhecimento de padrões. Devido à capacidade de aprendizagem a RNA tem sido utilizada no desenvolvimento de ferramentas que auxiliem tarefas complexas que depende muito da experiência do operador.

1.3.1 Modelo do Neurônio

O neurônio é uma unidade de processamento de informação fundamental para a operação de uma rede neural. Na Figura 15 é mostrado o modelo de um neurônio, que forma a base para a modelagem artificial de Redes Neurais.

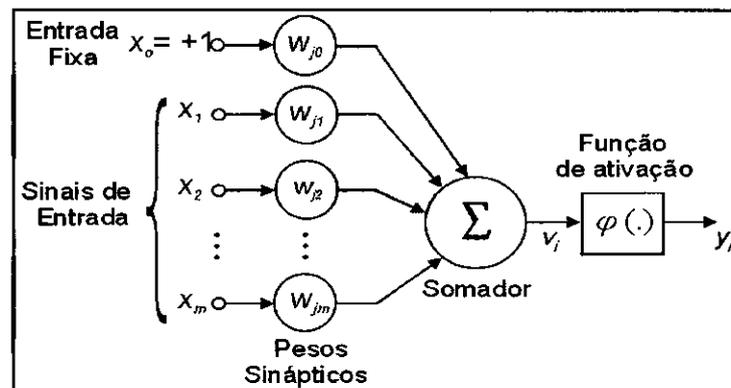


Figura 15 - Modelo de um neurônio artificial.

Como mostrado na Figura 15, pode-se identificar três elementos básicos no modelo neural:

- O conjunto de sinapses ou conexões, onde cada uma tem um peso sináptico específico.
- O sinal x_j na entrada da sinapse conectada ao neurônio k é multiplicado pelo peso sináptico w_{jk} .
- O somador é um operador utilizado para somar os sinais de entrada ponderados pelos respectivos pesos sinápticos, em cada ligação. O somador é na realidade uma combinação linear das entradas do neurônio.
- A função de ativação serve para limitar a saída de um neurônio. Tipicamente, o intervalo de amplitude normalizado da saída é compreendido entre o intervalo $[0,1]$ ou alternativamente, $[-1,1]$.

O modelo do neurônio também inclui um bloco chamado de polarização b_k (*bias*), aplicado externamente, que tem o efeito de incrementar ou decrementar a entrada da função de ativação, dependendo do seu valor ser positivo ou negativo, respectivamente.

Descreve-se um neurônio k , escrevendo o seguinte par de equações:

$$v_k = \sum_{j=0}^m w_{kj} x_j ; \quad (12)$$

$$y_k = \varphi(v_k). \quad (13)$$

Sendo x_0, x_1, \dots, x_n os sinais de entrada, com x_0 assumindo um valor constante igual a 1; $w_{k0}, w_{k1}, \dots, w_{kn}$ os pesos sinápticos do neurônio k , com w_{k0} sendo igual a b_k que a representação a polarização ou bias; v_k conhecido como potencial de campo induzido do neurônio ou potencial de ativação; $\varphi(\cdot)$ a função de ativação do neurônio; e y_k o sinal de saída do neurônio.

Dai, podem-se escrever as Equações (12) e (13) da seguinte forma:

$$v_k = \sum_{j=0}^m w_{kj} x_j + b_k ; \quad (14)$$

$$y = \varphi(v_k + b_k), \quad (15)$$

Dependendo do valor de b_k ser positivo ou negativo, a relação entre y_k e a saída do somador v_k é modificada.

1.3.2 Tipos de Função de Ativação

A função de ativação, identificada por $\varphi(v_k)$, define a saída de um neurônio k em função de um campo local induzido v_k . Os três tipos básicos de função de ativação são: limiar brusco, limiar lógico e sigmóide.

Função Limiar Brusco - Também conhecida como *Heaviside*. Este tipo de função a saída $\varphi(v_k)$ é mostrada pelo gráfico da Figura 16 e representada pela Equação (16):

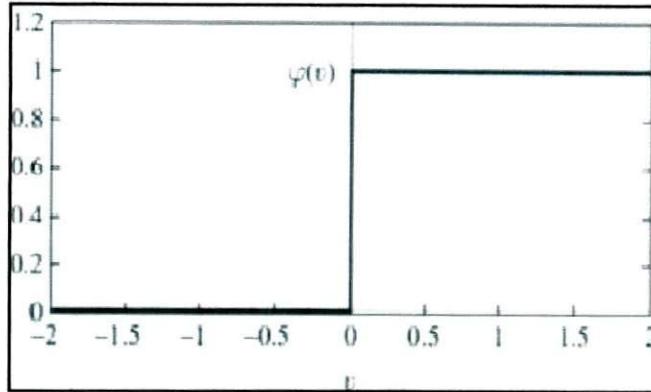


Figura 16 - Função de limiar brusco.

$$\varphi(v) = \begin{cases} 1 & \text{se } v \geq 0 \\ 0 & \text{se } v < 0 \end{cases}, \quad (16)$$

sendo v_k o campo local induzido do neurônio da Equação 14.

Função de Limiar Lógico – A função tem o seu comportamento representado na Figura 17 e se comporta como indicado pela Equação (17):

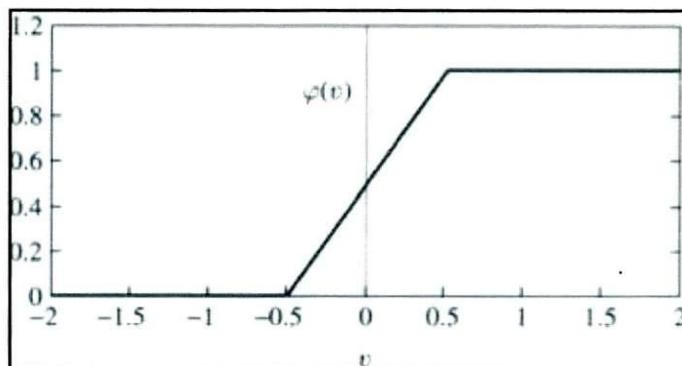


Figura 17 - Função limiar lógico.

$$\varphi(v_k) = \begin{cases} 1 & \text{se } v \geq \frac{1}{2} \\ v & \text{se } -\frac{1}{2} < v < \frac{1}{2} \\ 0 & \text{se } v \leq -\frac{1}{2} \end{cases}. \quad (17)$$

Função Sigmóide - A função sigmóide mostrada na Figura 18 tem um formato de s, ela é definida como uma função estritamente crescente que exibe um bom comportamento

intermediário entre o modelo linear e o não-linear. Um dos exemplos de função sigmóide é a função logística Equação (18),

$$\varphi(v) = \frac{1}{1 + \exp(-av)} , \quad (18)$$

em que, a é o parâmetro da função. Quando o valor de a é alterado consegue-se obter funções com inclinação diferentes.

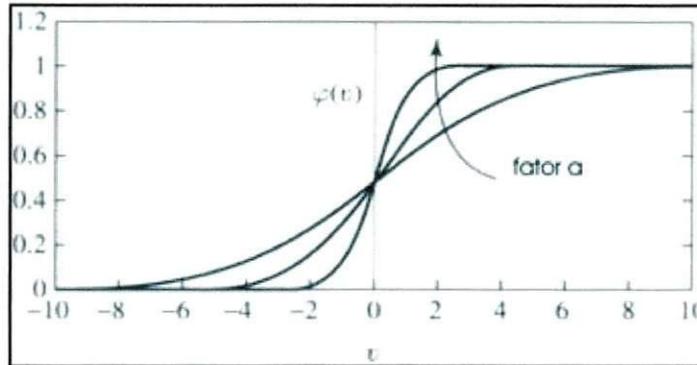


Figura 18 - Função de sigmóide.

Existem ainda funções de ativação sigmóide que assumem valores entre -1 e 1 como a função tangente hiperbólica, seu gráfico a anti-simétrico em relação à origem.

1.3.3 Arquitetura de Redes Neurais

Basicamente, a estrutura de uma rede neural está ligada à forma como ela será treinada, pois existem vários algoritmos de aprendizagem. Em geral, é possível identificar três classes de arquitetura de rede diferentes.

1.3.3.1 Redes de Única Camada

Na arquitetura de Rede de Única Camada tem-se uma camada de entrada de nós de fonte, que se projeta sobre uma camada de saída de neurônio. Na Figura 19 é representada uma estrutura de uma rede de única camada.

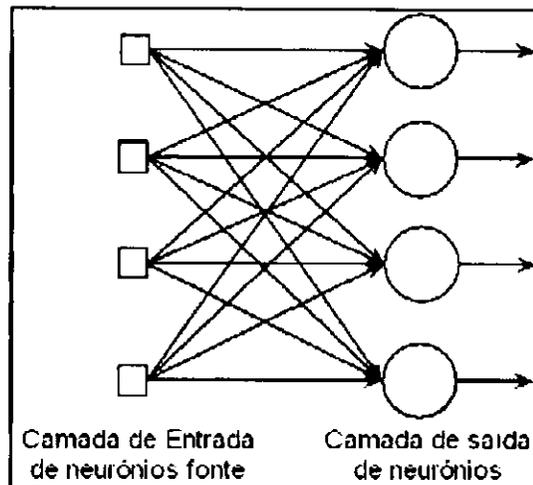


Figura 19 - Arquitetura de uma rede de camada única.

1.3.3.2 Rede de Múltiplas Camadas

A Rede de Múltiplas Camadas apresenta camadas intermediárias ou escondidas, compostas de nós chamados de unidades escondidas ou neurônios escondidos. Neste caso, as unidades que pertencem à camada de entrada são conhecidas como unidades sensórias. Pela adição de uma ou mais camadas escondidas, a rede é capaz de extrair mais informação sobre os dados de entrada. O sinal de entrada se propaga através da rede, camada por camada, até que uma resposta seja produzida na saída da rede. Na Figura 20 é possível visualizar a estrutura de uma rede neural de múltiplas camadas.

Os sinais de saída de uma camada são utilizados na próxima camada, e assim por diante até o fim da rede. A rede neural é chamada de totalmente conectada, quando cada um dos nós de uma camada de rede está conectado a todos os nós da camada adjacente seguinte. Porém, se alguns dos elos de comunicação (conexões sinápticas) não estiverem conectados, a rede recebe o nome de parcialmente conectada.

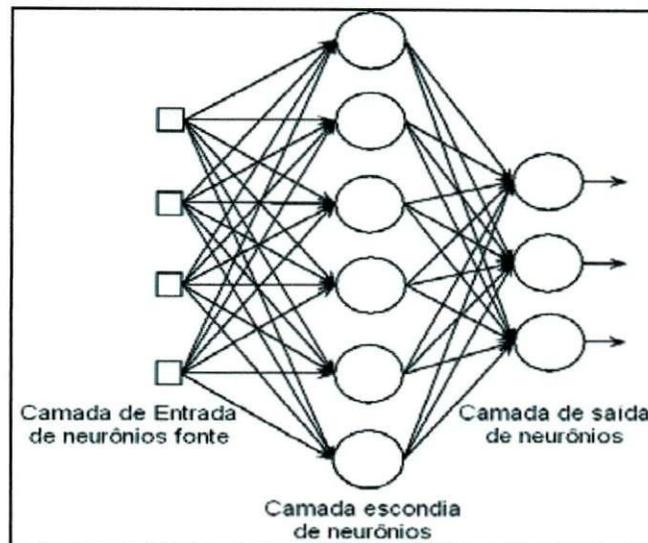


Figura 20 - Arquitetura de uma rede de múltiplas camadas.

1.3.3 Redes Recorrentes

Na arquitetura de redes recorrentes existe um laço de realimentação onde o sinal de saída volta para a entrada de todos os outros neurônios. A presença de laço de realimentação tem influência na capacidade de aprendizagem da rede e no seu desempenho. Um laço de realimentação envolve elementos de atraso unitário (representado por z^{-1}), o que ocasiona um comportamento dinâmico não-linear da rede, como indicado pela Figura 21.

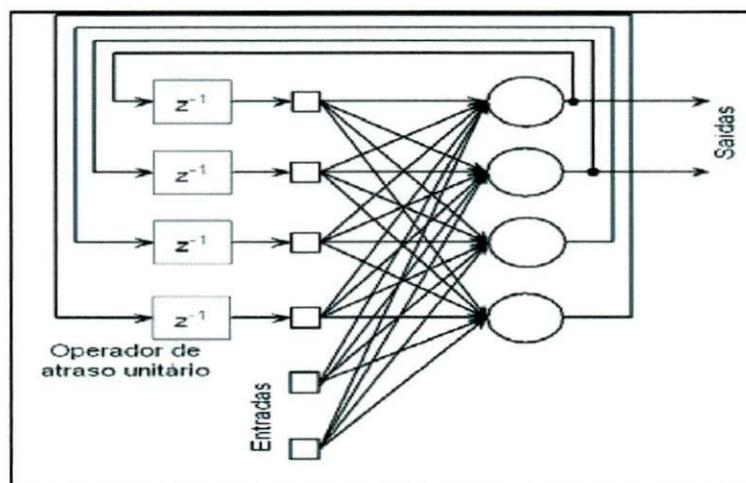


Figura 21 - Arquitetura de uma rede recorrente de camada única.

1.3.4 Processo de Aprendizagem

O processo de aprendizagem é a sua habilidade de aprender utilizando o seu ambiente e de aperfeiçoar o seu próprio desempenho. A rede neural aprende através de um processo iterativo de ajustes aplicados a seus pesos sinápticos e níveis de bias.

Segundo (SIMON, 2001), a aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros ocorre. O processo de aprendizagem obedece a uma seqüência de eventos:

- A rede neural é estimulada por uma entrada;
- A rede neural efetua modificação em seus parâmetros livres após a estimulação;
- A rede neural começa a responder de modo diferente, a partir das modificações efetuadas.

O conjunto de regras utilizado para efetuar a aprendizagem da rede neural é chamado algoritmo de aprendizagem. Existem inúmeras formas de algoritmo de aprendizagem onde cada uma contém características diferentes. Basicamente, os algoritmos diferem na forma como é formulado o ajuste de pesos sinápticos de um neurônio.

1.3.5 Redes Perceptron de Múltiplas Camadas

As redes *perceptron* de múltiplas camadas consistem em um conjunto de unidades sensoriais que constituem a camada de entrada, uma ou mais camadas ocultas e uma camada de saída de nós computacionais, em que o sinal de entrada se propaga para frente através da rede, camada por camada.

Para aplicação das redes neurais na solução de uma diversidade de problemas torna-se necessária à etapa de treinamento da rede, onde ela poderá adquirir o conhecimento sobre a aplicação. O principal algoritmo de treinamento de redes de múltipla camada é conhecido como algoritmo de retropropagação de erro, do inglês *error backpropagation*. Ele consiste de um tipo de treinamento supervisionado, sendo necessário informar o conjunto de sinais de saída desejados para a rede, para que a cada

sinal de entrada apresentado, a rede tenta corrigir o erro entre seu sinal de saída calculado e o desejado.

Uma rede perceptron de múltiplas camadas tem três características distintivas:

1. No modelo de cada neurônio é incluída uma função de ativação não-linear suave (diferenciável em qualquer ponto). Uma forma utilizada de não-linearidade é a função logística,

$$\varphi(v) = \frac{1}{1 + \exp(-v)},$$

em que, v é o campo local induzido do neurônio, $\varphi(v)$ é a saída do neurônio.

2. A rede contém uma ou mais camadas de neurônio ocultos. Os neurônios ocultos dotam a rede da capacidade de aprender tarefas complexas extraíndo as características mais significativas dos padrões de entrada.
3. A rede exibe um alto grau de conectividade, determinado pelas sinapses da rede. Uma modificação na conectividade da rede requer uma mudança na população das conexões sinápticas ou de pesos.

1.3.6 Treinamento da Rede

A aprendizagem de uma rede consiste em conseguir efetuar adaptação de seus pesos, de modo que a rede fique bem treinada e aprenda o suficiente. Desta forma, o processo de aprendizagem se transforma em uma escolha de parâmetros da rede para um conjunto de dados. Neste caso, a escolha dos padrões a serem apresentados à rede durante a etapa de treinamento é de extrema importância. Eles devem ser representativos, de forma que seu conteúdo de informação seja o maior possível para a tarefa considerada.

No processo é comum tornar aleatória a ordem em que os exemplos são apresentados à rede de uma época para a seguinte. A aleatoriedade garante que os exemplos sucessivos apresentados à rede raramente pertençam à mesma classe e sim que eles sejam radicalmente diferentes.

Uma preocupação durante a etapa de treinamento consiste em evitar o treinamento excessivo da rede de forma que o reajuste efetuado nos pesos ao invés de reduzir o erro gradativamente, passe a aumentar a cada iteração, chamado de *overtraining* (treinamento excessivo). Ele pode fazer com que o algoritmo nunca venha a convergir e que nenhum critério de parada do algoritmo seja atingido. Nesse ponto,

diz-se que a rede não consegue adquirir a capacidade de generalização. O treinamento excessivo pode ser causado pela arquitetura da rede ser inadequada para aquela aplicação ou o conjunto de treinamento que não consiga retratar de forma representativa o conhecimento que deve ser adquirido pela rede.

Para tentar evitar os problemas citados é possível utilizar a validação cruzada, que consiste em um banco de treinamento dividido em três conjuntos: treinamento, validação e teste. Os conjuntos são utilizados durante o processo de treinamento e servem para monitorar a evolução do erro médio quadrático na saída da rede com o intuito de minimizar treinamento excessivo.

O conjunto de treinamento é utilizado para escolher o modelo da rede que melhor se adapte, enquanto que o conjunto de validação é utilizado apenas para validar o modelo, através da apresentação de padrões completamente distintos dos utilizados durante o treinamento. O conjunto de teste é utilizado para testar o desempenho da rede na tarefa considerada. O conjunto de validação é aplicado à rede periodicamente a fim de se calcular o erro médio quadrático no conjunto. Os pesos das ligações da rede se mantêm inalterados, servindo para o monitoramento da capacidade da rede de generalizar, mas também podendo ser utilizado como critério de parada.

1.3.7 Algoritmo de Retropropagação

A aprendizagem por retropropagação de erro consiste em duas etapas:

Propagação: O vetor de entrada é aplicado aos nós sensoriais da rede e seu efeito se propaga na rede, camada por camada, produzindo um vetor de saída real da rede;

Retropropagação: Os pesos sinápticos são todos ajustados de acordo com uma regra de correção de erro. É calculado o sinal de erro que é propagado para trás através da rede onde o sinal de saída é utilizado no reajuste dos pesos sinápticos.

O sinal de erro “ e ” na saída do neurônio j é calculado na n -ésima iteração da seguinte forma:

$$e_j(n) = d_j(n) - y_j(n), \quad (19)$$

sendo y_j , a saída do neurônio e d_j , o valor esperado para saída.

O valor instantâneo de erro é definido para um neurônio j como:

$$\frac{1}{2} e_j^2(n), \quad (20)$$

em cada iteração e o valor instantâneo da energia total de erro $\xi(n)$ é obtido somando-se

os termos $\frac{1}{2}e_j^2(n)$ de todos os neurônios da camada de saída.

$$\xi(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n), \quad (21)$$

em que, C indica o conjunto de todos os neurônios da camada de saída da rede.

A energia média $\xi(n)_{med}$ é calculada a partir da energia instantânea:

$$\xi(n)_{med} = \frac{1}{N} \sum_{n=1}^N \xi(n), \quad (22)$$

sendo N o número total de padrões contidos num conjunto de treinamento.

Tanto a energia instantânea quanto à energia média são funções de todos os parâmetros livres da rede. Para um dado conjunto de treinamento, $\xi(n)_{med}$ representa a função de custo como uma medida do desempenho da aprendizagem. O objetivo do processo de aprendizagem é ajustar os parâmetros livres da rede para minimizar $\xi(n)_{med}$.

Um método simples de realizar o treinamento da rede consiste em apresentar o conjunto de treinamento de padrão em padrão até formar uma época, apresentação completa do conjunto ou de um subconjunto de treinamento que está sendo processado. Os ajustes dos pesos são realizados de acordo com o valor do erro calculado para cada padrão apresentado a rede. A média aritmética das alterações individuais de peso sobre o conjunto de treinamento é uma estimativa da alteração dos pesos.

O ajuste dos pesos pode ser efetuado de duas maneiras: seqüencial ou por lote. No modo seqüencial os pesos são atualizados de padrão em padrão até completar toda a apresentação do conjunto de treinamento, com o ajuste dos pesos sendo realizado de acordo com respectivos erros calculados para cada padrão apresentado à rede. Enquanto que no reajuste por lote, o valor do reajuste dos pesos também é calculado para cada padrão, mas os pesos não são atualizados a cada passo. Na realidade, é feita uma estimativa da alteração real que resultaria na minimização da função de custo ξ_{med} sobre todo o conjunto de treinamento, que consiste em atualizar os pesos apenas ao final de cada apresentação do conjunto de treinamento, utilizando como parâmetro de reajuste a média aritmética das alterações individuais de cada peso calculadas para todo conjunto de treinamento.

O algoritmo de retropropagação tem o intuito de minimizar a energia do erro, ou função de erro $\xi(n)$ aplicando uma correção de $\Delta w_{ji}(n)$ ao peso w_{ji} que é proporcional à derivada parcial $\frac{\partial \xi}{\partial w_{ji}}$. Pela regra da cadeia:

$$\frac{\partial \xi(n)}{\partial w_{ji}(n)} = \frac{\partial \xi(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)}, \quad (23)$$

a derivada parcial $\frac{\partial \xi}{\partial w_{ji}}$ representa um fator de sensibilidade, determinando a direção de

busca no espaço de pesos, para o peso sináptico w_{ji} .

Considerando as equações,

$$v_j = \sum_{i=0}^m w_{ij}(n) x_j(n),$$

$$y_j(n) = \varphi_j(v_j(n)),$$

$$\xi(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n),$$

$$e_j(n) = d_j(n) - y_j(n)$$

e realizando as diferenciações, encontram-se os resultados abaixo:

$$\frac{\partial \xi(n)}{\partial e_j(n)} = e_j(n) \quad (24)$$

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1 \quad (25)$$

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi_j'(v_j(n)), \quad (26)$$

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = x_j(n) \quad (27)$$

Agora substituindo as equações (24), (25), (26) e (27) em (23) obtém-se:

$$\frac{\partial \xi(n)}{\partial w_{ji}(n)} = -e_j(n) \varphi_j'(v_j(n)) x_j(n), \quad (28)$$

o sinal negativo da formula indica o sentido decrescente do gradiente no espaço de pesos. A correção $\Delta w_{ji}(n)$ aplicada a $w_{ji}(n)$ é definida pela regra do delta:

$$\Delta w_{ji}(n) = \eta \frac{\partial \zeta(n)}{\partial w_{ji}(n)}, \quad (29)$$

sendo η , o parâmetro da taxa de aprendizagem do algoritmo de retropropagação, utilizando (29) e (30) encontra-se:

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n), \quad (30)$$

que é o cálculo para o neurônio j da camada de saída, em que o termo $\delta_j(n)$ é chamado de gradiente local. O cálculo do gradiente local para o neurônio na última camada é encontrado da seguinte forma:

$$\begin{aligned} \delta_j(n) &= -\frac{\partial \zeta(n)}{\partial v_j(n)}, \\ &= -\frac{\partial \zeta(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)}, \\ &= e_j(n) \varphi'_j(v_j(n)). \end{aligned} \quad (31)$$

Para um neurônio j que está em uma camada intermediária não é possível saber o valor da saída do neurônio, logo o valor de $\delta_j(n)$ é encontrado da seguinte forma:

$$\delta_j(n) = -\frac{\partial \zeta(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} = -\frac{\partial \zeta(n)}{\partial y_j(n)} \varphi'_j(v_j(n)) \quad (32)$$

A partir da função de erro para a saída de um neurônio j , a Equação (21) pode-se escrever como:

$$\frac{\partial \zeta(n)}{\partial y_j(n)} = \sum_{k \in D} e_k(n) \frac{\partial e_k(n)}{\partial y_j(n)}, \quad (33)$$

sendo D o conjunto de neurônio da camada intermediária em questão. Aplicando a regra da cadeia na Equação (34), tem-se:

$$\frac{\partial \zeta(n)}{\partial y_j(n)} = \sum_{k \in D} e_k(n) \frac{\partial e_k(n)}{\partial y_k(n)} \frac{\partial v_k(n)}{\partial y_j(n)}, \quad (34)$$

Como o erro no neurônio k pode ser escrito por:

$$e_k(n) = d_k(n) - y_k(n) = d_k(n) - \varphi'_k(v_k(n)), \quad (35)$$

e derivando a Equação (35) com relação a $v_k(n)$, obtêm-se,

$$\frac{\partial e_k(n)}{\partial v_k(n)} = -\varphi'_k(v_k(n)), \quad (36)$$

Pode-se dizer também que o valor de ativação $v_k(n)$ para o neurônio k é dado por:

$$v_k = \sum_{i=0}^m w_{k_j}(n) y_k(n), \quad (37)$$

sendo m o número total de entradas aplicadas ao neurônio k , onde y_0 é fixado em 1 e w_{k0} o peso sináptico para a entrada *bias*.

Derivando a equação (38) parcialmente com relação a $y_j(n)$, tem-se:

$$\frac{\partial v_k(n)}{\partial y_j(n)} = w_{k_j}(n), \quad (38)$$

Das equações (36) e (38), pode-se escrever a Equação (33), como:

$$\frac{\partial \zeta(n)}{\partial y_j(n)} = -\sum_{k \in D} e_k(n) \varphi'_k(v_k(n)) w_{k_j}(n) = -\sum_{k \in D} \delta_k(n) w_{k_j}(n), \quad (39)$$

finalmente substituindo a Equação (40) na Equação (32) pode-se escrever o gradiente local para o neurônio j localizado em uma camada intermediária como:

$$\delta_j(n) = \varphi'_j(v_j(n)) \sum_j \delta_k(n) w_{k_j}(n) \quad (40)$$

1.3.8 Taxa de Aprendizagem

No algoritmo de retropropagação quanto menor for a taxa de aprendizagem η menor será as variações nos pesos sináptico da rede, causando uma aprendizagem lenta. Um outro ponto a ser analisado é quando a taxa de aprendizagem η for muito grande. Para acelerar a aprendizagem, grandes modificações nos pesos sinápticos podem provocar instabilidade.

1.3.9 Critério de Parada

Existem alguns critérios razoáveis de parada onde cada um tem suas qualidades e que podem ser usadas para encerrar o ajuste de pesos. Para se formular um critério de parada é necessário analisar as propriedades de mínimo local ou global alcançado na superfície de erro. Uma condição necessária para que um mínimo seja alcançado é que o

vetor gradiente seja muito pequeno. Outra propriedade utilizada para formular critério de parada é o fator da função custo ou média de erro $\xi_{med}(w)$, onde se utiliza a taxa absoluta de variação do erro médio quadrático por época como critério de parada do treinamento. Pode ser utilizado como critério de parada o número de épocas durante o treinamento, também como o valor do erro do conjunto de validação.

1.3.10 Algoritmo Resilient Propagation

Durante o processo de aprendizagem a rede neural pode não conseguir encontrar o ajuste de pesos que atenda as necessidades da implementação, o problema pode estar relacionado com diversos fatores com qualidade dos dados, modo de processamento dos dados, configuração de rede ou tipo de algoritmo de treinamento, que devido as suas características não consegue generalizar a solução do fenômeno avaliado. Para tentar solucionar o problema referente ao tipo de algoritmo foram desenvolvidos alguns modelos que utilizam como mecanismo de otimização a adaptação dos parâmetros da rede durante a fase de aprendizagem.

O algoritmo de retropropagação de erros, apesar de apresentar simplicidade na implantação, possui uma convergência muito lenta e pode chegar até a não convergir, mesmo com a utilização de algumas heurísticas para melhorar o desempenho do método. Isso se deve ao fato de que o método opera segundo uma aproximação linear da função de erro da rede na vizinhança local do ponto de operação, descrito pela matriz de pesos da rede, (SIMON, 1999).

Um fator que contribui para a não convergência está relacionado com o fato das atualizações dos pesos não depender somente do valor da taxa de aprendizagem, mas também do valor da derivada parcial do erro. Para tentar solucionar a dependência foi criado um algoritmo chamado *resilient propagation* (Rprop), que consiste em uma variação do algoritmo backpropagation, onde a adaptação dos pesos é realizada fazendo uso do sinal da derivada do erro.

No Rprop, os pesos e a taxa de aprendizagem são alterados apenas após a apresentação de todo o conjunto de treinamento, reajuste por lote (*batch*). Cada peso $w_{j,i}(n)$ possui sua própria taxa de variação $\Delta w_{j,i}(n)$, a qual varia em função do número de iterações, como mostrado na Equação (41) (SILVA, 2004).

$$\Delta_{ji}(t) = \begin{cases} \eta^+ \cdot \Delta_{ji}(t-1), & \text{se } \frac{\partial \xi}{\partial w_{ji}}(t-1) \cdot \frac{\partial \xi}{\partial w_{ji}}(t) > 0 \\ \eta^- \cdot \Delta_{ji}(t-1), & \text{se } \frac{\partial \xi}{\partial w_{ji}}(t-1) \cdot \frac{\partial \xi}{\partial w_{ji}}(t) < 0 \\ \Delta_{ji}(t-1), & \text{caso contrário} \end{cases} \quad (41)$$

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

A Equação (41) mostra que a mudança do sinal das derivadas parcial indica que ocorreu um salto sobre um ponto de mínimo da função $\xi(t)$, logo o valor do reajuste deve ser diminuído. Por outro lado derivadas com mesmo sinal indicam que o sistema está se movendo permanentemente em uma única direção. O reajuste dos pesos sinápticos é encontrado pela Equação (42):

$$\Delta w_{ji}(t) = \begin{cases} -\Delta_{ji}(t), & \text{se } \frac{\partial \xi}{\partial w_{ji}}(t) > 0 \\ +\Delta_{ji}(t), & \text{se } \frac{\partial \xi}{\partial w_{ji}}(t) < 0 \\ 0, & \text{caso contrário} \end{cases} \quad (42)$$

Como mostrado na Equação (42) o reajuste dos pesos sinápticos depende apenas do sinal da derivada de $\xi(t)$. Devido às características do RPROP sua convergência é mais rápida e apresenta poucas variações com relação aos parâmetros de treinamento.

2 Materiais e Métodos

Os sinais DPs foram adquiridos em ensaios elétricos utilizando metodologia descrita em (NERI, 2008). Os ensaios elétricos foram realizados no Laboratório de Alta Tensão da Universidade Federal de Campina Grande. Com o intuito de gerar sinais de corona com ponta em alta e com ponta aterrada foi utilizado um arranjo experimental de um kit didático de alta tensão fabricado pela Messwandler-Bau®.

Os sinais de descargas parciais foram medidos em transformadores de potencial e de distribuição da classe de tensão de 15 kV, que apresentavam descargas parciais. O sinal cru, ou seja, sem tratamento foi medido em uma saída auxiliar do sistema comercial de medição e análise de descargas parciais fabricado pela Lemke, a banda passante do sistema comercial é de um megahertz. Os sinais foram medidos e registrados por um osciloscópio digital com banda passante de até um gigahertz.

A natureza das descargas medidas foi assegurada pelo diagnóstico computadorizado do equipamento comercial, o resultado foi agrupado da seguinte forma:

- 1) Descargas Parciais no TP.
- 2) Descargas Parciais com Corona (ponta em alta).
- 3) Corona (ponta em alta).
- 4) Descargas Parciais com Corona (ponta em baixa).
- 5) Corona (ponta em baixa).

Na Figura 22 é possível observar o transformador de alta tensão, capacitor de acoplamento e a bobina de bloqueio utilizados no arranjo experimental para medição dos sinais de descargas parciais e corona.



Figura 22 - Arranjo do sistema de medição de descargas parciais com capacitor de acoplamento e bobina de bloqueio (NERI, 2008).

Os sinais de descargas foram processados aplicando operadores estáticos na extração de características com relação ao formato da distribuição dos sinais de DPs. A técnica de processamento utilizou uma configuração que dividia o sinal em janela com relação ao ângulo da senóide seguindo metodologia apresentada em (GULSKI, 1993), como mostrado na Figura 23. Foram realizados dois tipos de processamento com o sinal contido nas janelas, no primeiro foi encontrado o valor máximo das DPs e no segundo foi calculado o valor médio das DPs em cada janela. Com as informações de cada janela foram calculados os operadores estatísticos dos sinais.

Como os operadores assimetria e curtose avaliam a forma como as descargas estão dispostas em relação à distribuição normal, o cálculo das médias para cada janela pode modificar as características da descarga e assim influenciar o resultado. Este problema foi contornado com a utilização de 3600 de janelas, diminuindo assim o intervalo de tempo de cada uma tornando o sinal mais fiel.

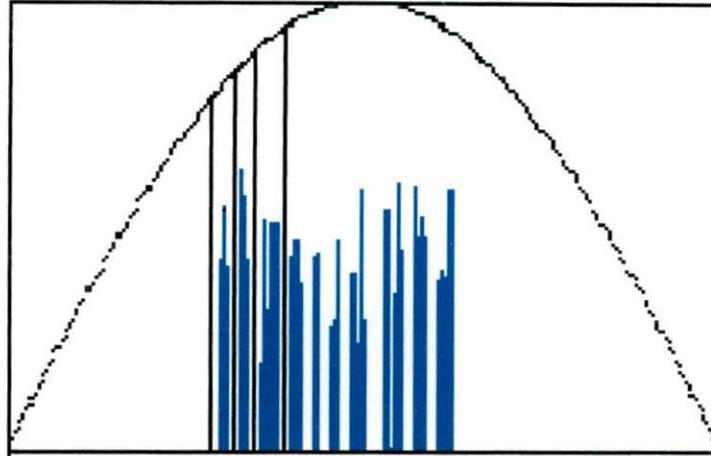


Figura 23 - Semi-ciclo do sinal de DP dividido em janelas.

Ao final da etapa de processamento, utilizando uma rotina escrita em Matlab 7.0 © (MATHWORKS, 2004), obteve-se uma base de dados com 1250 padrões de DPs contendo 250 padrões para cada tipo de DPs constituídos de 18 características cada sinal como apresentado na seqüência abaixo:

- 1) Correlação Cruzada;
- 2) Assimetria da descarga;
- 3) Desvio Padrão no semi-ciclo positivo (Médias);
- 4) Desvio Padrão no semi-ciclo negativo (Médias);
- 5) Desvio Padrão no semi-ciclo positivo (Máximos);
- 6) Desvio Padrão no semi-ciclo negativo (Máximos);
- 7) Assimetria no semi-ciclo positivo (Médias);
- 8) Assimetria no semi-ciclo negativo (Médias);
- 9) Assimetria no semi-ciclo positivo (Máximos);
- 10) Assimetria no semi-ciclo negativo (Máximos);
- 11) Curtose no semi-ciclo positivo (Médias);
- 12) Curtose no semi-ciclo negativo (Médias);
- 13) Curtose no semi-ciclo positivo (Máximos);
- 14) Curtose no semi-ciclo negativo (Máximos);
- 15) Centro da distribuição no semi-ciclo positivo;
- 16) Centro da distribuição no semi-ciclo negativo;
- 17) Tensão inicial da descarga no semi-ciclo positivo;
- 18) Tensão inicial da descarga no semi-ciclo negativo;

Os dados foram organizados em uma base de dados onde cada sinal recebeu um código binário de três bits para identificar a sua origem como mostrado na Tabela 1. As 18 características foram normalizadas com a finalidade de adapta os dados para o treinamento da RNA. Na Figura 24 é apresentado o arquivo da base de dados.

Tabela 1 - Código binário de cada tipo de DP.

Origem da Descarga	Código
Descargas Parciais	011
Descargas Parciais com Corona (ponta em alta)	101
Corona (ponta em alta)	110
Descargas Parciais com Corona (ponta em baixa)	111
Corona (ponta em baixa)	100

	1	2	3	4	5	6	7	8
1	-0.61509	-0.6986	-0.66084	-0.67917	-0.67627	-0.68694	-0.63047	-0.6986
2	-0.80547	-1.0038	1.7694	3.1209	-0.1031	0.27822	-0.8043	-1.0038
3	2.97	-0.78941	-0.71816	-0.77841	-0.66879	-0.60872	3.2099	-0.78941
4	-1.3249	0.048086	0.076438	-0.15176	-1.1194	-0.6933	-1.3432	-0.031806
5	2.9804	-0.84269	-0.67941	-0.81168	-0.64104	-0.6424	3.1516	-0.84269
6	-1.3086	0.081216	0.17173	-0.11391	-1.0847	-0.81376	-1.2898	-0.024068
7	-0.50872	-0.83916	0.11868	2.8726	1.0528	1.7855	-0.48837	-0.83916
8	1.959	0.044684	0.093882	0.082194	0.15341	2.0257	2.1832	0.068719
9	-0.57503	0.23802	-0.16715	2.0724	-0.0027705	1.3576	-0.55963	0.23802
10	1.9795	0.11986	0.16438	0.099137	-0.21422	2.2396	1.979	0.1189
11	-0.49805	-0.53671	-0.23725	3.6361	0.83385	1.6325	-0.49398	-0.53671
12	1.89	-0.26819	-0.23469	-0.24599	-0.1358	2.04	2.4656	-0.25326
13	-0.535	0.4472	-0.4261	1.624	-0.2237	0.8621	-0.5322	0.4472
14	1.7934	-0.22552	-0.18693	-0.24382	-0.33986	2.4738	1.7925	-0.22541
15	0.69413	-2.2077	-0.098632	-0.26849	0.94828	-0.31737	0.66256	-2.2077
16	0.66059	0.43602	-0.10485	0.36121	0.15206	-0.78285	-0.57419	0.37734
17	0.78229	-1.2035	-1.2512	-1.2544	0.25689	0.27216	0.84626	-1.2239
18	-1.1817	1.1744	1.1494	1.1967	0.26844	0.37322	-1.1614	1.1676
19	1	1	1	1	1	1	1	1
20	0	1	0	1	0	0	0	1
21	0	0	1	0	1	1	0	0

Figura 24 - Parte do arquivo da base de dados.

3 Simulações e Resultados

O trabalho tem a finalidade de treinar uma estrutura de rede neural artificial para o diagnóstico de descargas parciais. O treinamento da RNA foi realizado com uma base de dados com o sinal de DP, que passaram por uma fase de processamento para retirada extração de característica que a representassem as formas da descarga, utilizando operadores estatísticos. Com a base de dados foram realizados testes com diferentes configurações de entrada (estímulos) na RNA para avaliar a influência das características extraída do sinal de DP na identificação de padrões.

Os treinamentos foram divididos em etapas para a avaliação de diferentes configurações de base de dados. Os testes foram realizados em um micro computador Athlon® 2.0 GHz com 1 gigabyte de memória RAM, com rotinas escritas em Matlab 7.0® utilizando recursos do pacote *toolbox neural network* (MATHWORKS, 2003).

3.1 Primeira etapa

A primeira etapa do processo de treinamento da RNA teve como objetivo avaliar a capacidade dos operadores estatísticos na caracterização dos cinco tipos de sinal de DPs existentes. Para isso foram adotadas como entradas da rede as 18 características obtidas no processamento do sinal. A base de dados foi dividida em 70% para treinamento e 30% para teste.

Os critérios de parada do treinamento adotados foram o número de épocas igual a 1000, erro médio quadrático 0 e valor do gradiente da ordem 10^{-10} . A estrutura de rede adotada foi uma *perceptron* de múltiplas camadas, treinada com algoritmo de treinamento Rprop e pesos sinápticos foram iniciados aleatoriamente entre no intervalo de 0 a 1.

A escolha do número de neurônios da camada intermediária foi realizada através de teste onde o desempenho da rede foi monitorado, mediante os resultados obtidos foram efetuados aumentos gradativos do número de neurônios. A estrutura de rede que obteve melhor desempenho na fase de teste com 97,9% de acerto, tinha 25 neurônios na camada oculta. Suas funções de ativação foram tangente hiperbólica na camada oculta e logística na camada de saída. Na Figura 25 é apresentado o esquema da estrutura de rede utilizada durante a fase de aprendizagem, a camada de entrada é composta por 18

neurônios e camada saída da rede por três, que tem função de informar o código binário que identifica as DPs.

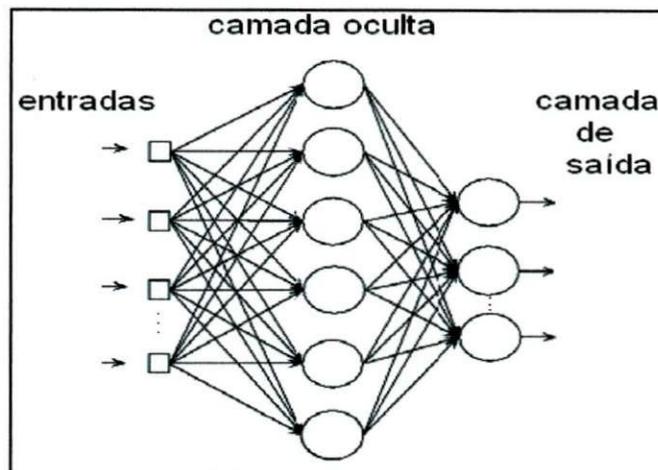


Figura 25 - Estrutura de RNA utilizada no treinamento.

O desempenho do erro médio quadrático na fase de treinamento é mostrado na Figura 26. Durante o treinamento o erro médio ficou estagnado em 10^{-3} a partir da época 450, o que caracterizou treinamento excessivo da estrutura. Porém o percentual de acerto foi satisfatório, relacionado à quantidade de informação apresentada na entrada da rede onde o grande número de características tornou a distinção dos padrões precisa.

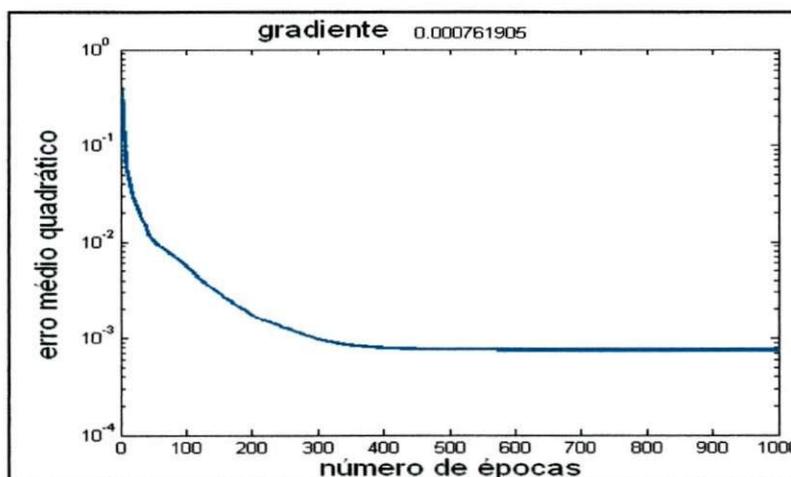


Figura 26 - Desempenho do erro primeira etapa sem validação.

A técnica de validação cruzada foi testada na mesma configuração de rede para avaliar o seu desempenho, seu objetivo principal evitar o treinamento desnecessário. A base de dados foi dividida em 50% para treinamento, 25% para validação e 25% para teste.

No teste com validação cruzada a parada do treinamento foi efetuada quando verificado o aumento do erro quadrático do conjunto de validação em relação ao erro do conjunto de treinamento 5 vezes consecutivas, neste caso a melhor configuração dos pesos sinápticos dentro do intervalo de 5 épocas foi adotada.

O resultado apresentado mostrou um número inferior de épocas durante o treinamento, porém um erro quadrático maior da ordem 10^{-2} , assegurado que a rede não foi treinada excessivamente. Na Figura 27 o desempenho do erro quadrático azul está diminuindo a cada época ao contrário do desempenho anterior e conseguiu acerto de 97,1%, em verde está o desempenho do erro do conjunto de validação cruzada.

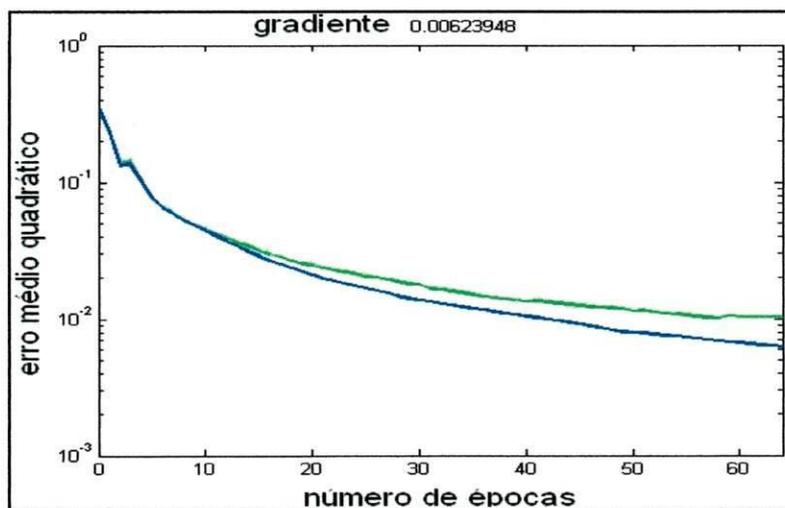


Figura 27 - Desempenho do erro primeira etapa com validação.

3.2 Segunda etapa

Confirmado o bom desempenho na classificação das descargas, um ponto necessário foi avaliar a influência dos dados de entradas em separado no percentual de acerto. Como os dados de entrada foram obtidos de duas formatações de sinal diferentes, o natural foi avaliar o desempenho da RNA para cada um deles. A primeira formatação utilizou os valores de máximos na avaliação do achatamento e espalhamento. A metodologia adotada é semelhante ao teste anterior, à estrutura de rede contém 25 neurônios na camada oculta modificando apenas o número de entradas. Abaixo é listada a seqüência de entradas da rede.

- 1) Correlação Cruzada;
- 2) Assimetria da descarga;

- 3) Desvio Padrão no semi-ciclo positivo (Máximos);
- 4) Desvio Padrão no semi-ciclo negativo (Máximos);
- 5) Assimetria no semi-ciclo positivo (Máximos);
- 6) Assimetria no semi-ciclo negativo (Máximos);
- 7) Curtose no semi-ciclo positivo (Máximos);
- 8) Curtose no semi-ciclo negativo (Máximos);
- 9) Centro da distribuição no semi-ciclo positivo;
- 10) Centro da distribuição no semi-ciclo negativo;
- 11) Tensão de inicial da descarga no semi-ciclo positivo;
- 12) Tensão de inicial da descarga no semi-ciclo negativo;

O resultado encontrado foi satisfatório 96,8% acerto e erro médio em torno de 10^{-3} . O percentual de acerto indica que é possível utilizando apenas os máximos do sinal em cada janela para classificar padrões das descargas, mesmo levando em conta à grande perda de informação que implica. Observando a Figura 28, verifica-se a presença de treinamento excessivo a partir da época 400.

Para melhorar o desempenho foi treinado uma estrutura utilizando validação cruzada. Como era de se esperar o resultado melhorou para 97,1%, acompanhado de uma redução no valor do erro quadrático devido à parada antecipada do treinamento, Figura 29. O acerto percentual teve o valor aumentado devido fato de a rede conservar sua melhor configuração após a interrupção do treinamento.

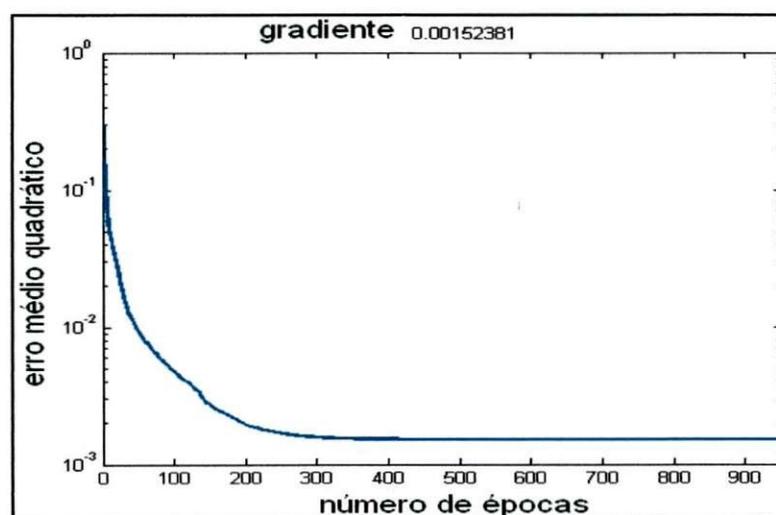


Figura 28 - Desempenho do erro segunda etapa sem validação (máximos).

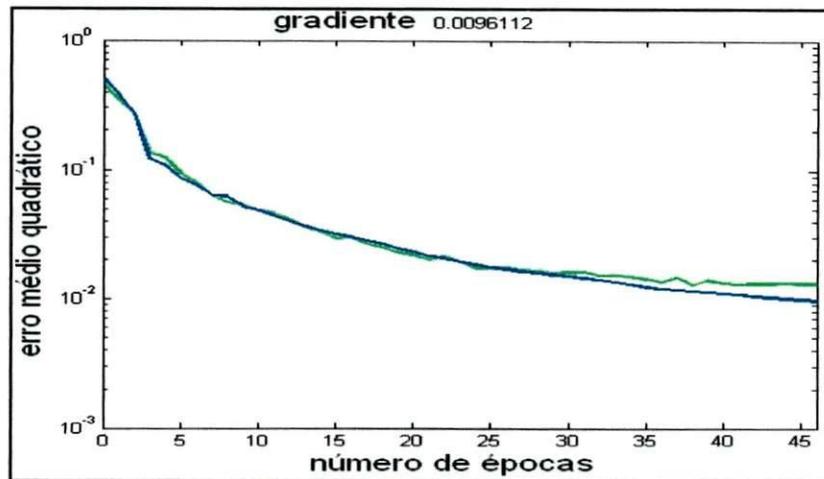


Figura 29 - Desempenho do erro segunda etapa com validação (máximos).

A outra configuração de base de dados continha os valores das médias das descargas em cada janela, com a finalidade de representar de forma mais fiel o valor da descarga durante o período de tempo de cada janela. Os dados de entrada da rede ficaram ordenados da seguinte forma:

- 1) Correlação Cruzada;
- 2) Assimetria da descarga;
- 3) Desvio Padrão no semi-ciclo positivo (Médias);
- 4) Desvio Padrão no semi-ciclo negativo (Médias);
- 5) Assimetria no semi-ciclo positivo (Médias);
- 6) Assimetria no semi-ciclo negativo (Médias);
- 7) Curtose no semi-ciclo positivo (Médias);
- 8) Curtose no semi-ciclo negativo (Médias);
- 9) Centro da distribuição no semi-ciclo positivo;
- 10) Centro da distribuição no semi-ciclo negativo;
- 11) Tensão de inicial da descarga no semi-ciclo positivo;
- 12) Tensão de inicial da descarga no semi-ciclo negativo.

O resultado foi satisfatório 95,2% de acerto, porém com treinamento excessivo a partir da época de número 500 mostrado na Figura 30. Com a validação o treinamento teve o desempenho de 96,8%, Figura 31.

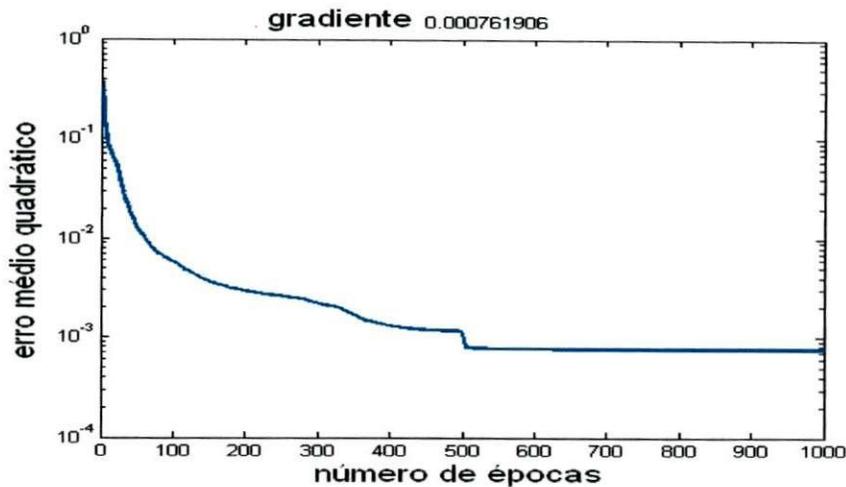


Figura 30 - Desempenho do erro segunda etapa sem validação (médias).

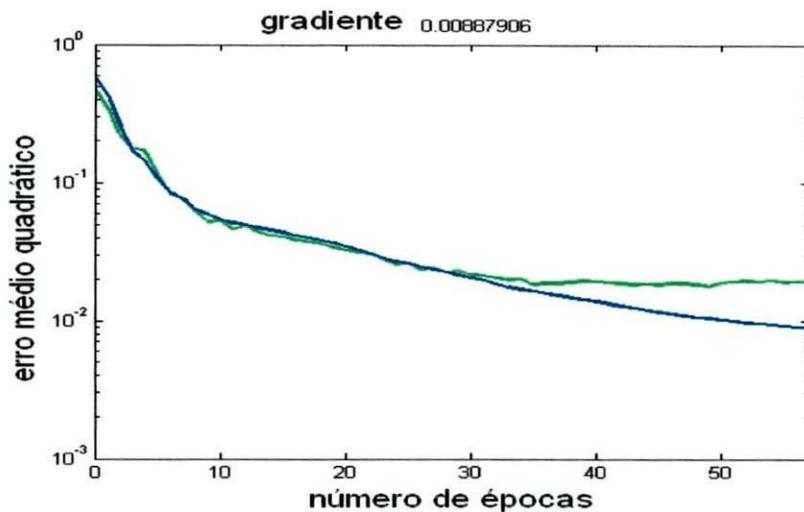


Figura 31 - Desempenho do erro segunda etapa com validação (médias).

O desempenho da RNA na fase de teste mostrou uma pequena diferença no percentual de acerto entre a de base de dado que utilizou os valores de máximo e as médias, ficando clara à viabilidade de ambas as configurações.

3.3 Terceira etapa

Uma configuração foi utilizada para investigar a influência dos operadores que quantifica a posição do conjunto de pontos em relação à sua média, sendo adotados os valores de máximo das janelas com base no desempenho dos testes anteriores. Outra configuração foi testada com intuito de avaliar a influência dos operadores que caracterizam as DPs pelas medidas de assimetria entre as contribuições do semi-ciclos

positivo e negativo. As entradas das duas configurações e os percentuais de acertos durante os testes são mostrados na Tabela 2.

Tabela 2 - Resultado dos testes da terceira etapa.

Parâmetros	Acerto sem Validação (%)	Acerto com Validação (%)
Correlação Cruzada	93,3	94,9
Assimetria da descarga entre os dois semi-ciclos		
Centro da distribuição no semi-ciclo positivo		
Centro da distribuição no semi-ciclo negativo		
Tensão de inicial da descarga no semi-ciclo positivo		
Tensão de inicial da descarga no semi-ciclo negativo		
Desvio Padrão no semi-ciclo positivo	92,8	93,6
Desvio Padrão no semi-ciclo negativo		
Assimetria no semi-ciclo positivo		
Assimetria no semi-ciclo negativo		
Curtose no semi-ciclo positivo		
Curtose no semi-ciclo negativo		

Como mostrado na literatura as descargas parciais podem ser caracterizadas pela sua posição em relação à senóide de tensão como também pela assimetria em cada semi-ciclo, servindo como uma impressão digital para cada tipo de situação. O percentual de acerto foi satisfatório e com valores bem próximos o que mostra uma equivalência na utilização de ambas as configurações para identificação de padrões.

3.4 Quarta etapa

Nesta etapa foram testadas redes tendo como entrada dois parâmetros, com a finalidade de avaliar individualmente o desempenho para cada um no resultado final do treinamento de uma RNA. Uma estrutura de rede com 6 neurônios na camada oculta foi treinada agora com apenas duas entradas tendo os mesmos critérios de parada e parâmetros dos testes anteriores. A idéia foi treinar uma mesma estrutura de rede com todas as possibilidades para verificar o grau de importância de cada um das características no resultado final. O resultado é apresentado na Tabela 3.

Tabela 3 - Resultado dos testes da quarta etapa.

Parâmetros	Acerto s/ Validação (%)	Acerto c/ Validação (%)
Correlação cruzada		
Assimetria das descargas	72,8	75,7
Desvio padrão (médias)	85,6	91,0
Desvio padrão (máximos)	86,9	86,9
Assimetria (médias)	83,7	85,6
Assimetria (máximos)	87,7	91,4
Curtose (médias)	83,5	77,0
Curtose (máximos)	85,6	86,9
Centro de distribuição	74,9	79,6
Tensão inicial de descarga	83,2	70,3

Os resultados apresentados mostram que na maioria dos casos o treinamento com validação cruzada tem desempenho melhor. O percentual de acerto onde a base de utilizou os valores de máximo foi maior que os treinados com valores de médias. Com relação à influência os operadores que quantificam a posição do conjunto de pontos em relação a uma distribuição de probabilidade (Assimetria, Curtose, Desvio) tiveram desempenho individuais superiores o que indica uma caracterização mais fiel do sinal.

3 Conclusões

Durante a avaliação das descargas utilizando redes neurais ficou comprovada a eficiência dos operadores estatísticos na caracterização do sinal, pois na primeira etapa de teste o percentual de acerto de 97,0% quando utilizado todos os operadores estatísticos.

Na segunda etapa quando foram testadas duas configurações diferentes uma utilizando máximo e outra com as médias dos valores nas janelas. Os percentuais de acerto tiveram valores aceitáveis e quase idênticos, fato motivado pela grande quantidade de janelas no sinal, cerca de 3600, o que impediu a perda significativa de informação durante o processamento.

A terceira etapa de teste evidencia que tanto o agrupamento dos operadores com finalidade de caracteriza a simetria do sinal ou a sua distribuição em relação à média tiveram percentual de acerto superior a 90,0%.

Na quarta etapa a avaliação individual dos operadores mostrou que o processamento do sinal utilizando os valores de máximo das janelas nos operadores Assimetria, Curtose e Desvio padrão tiveram melhor desempenho em relação aos demais.

Como esperado, em todos os testes realizados na maioria dos casos o melhor percentual de acerto foi atingido utilizando a técnica de validação cruzada que permitiu encontra uma configuração de rede neural evitando o treinamento excessivo. Os resultados dos testes mostraram que tanto as ferramentas de extração de características quanto à utilização de redes neurais artificiais são métodos confiáveis para classificação de sinais de DPs.

4 Bibliografia

- COSTA, E.G.; NEVES, W.L.A.: **Descargas Parciais**. Apostila, UFPB, 1996.
- GULSKI, E.; KREUGER, F.H.; **Recognition of discharge sources using statistical tools**, Proceedings of the 3rd International conference on Properties and Applications of Dielectric Materials, volume. 1, páginas. 39-42, Julho 1991.
- MATHWORK: **Toolbox Neural Networks 4.0**, Matlab Version 6.5, 2003.
- MATHWORK: **Matlab Version 7.0**, 2004.
- MOTA, H. O.: **Sistema de Aquisição e Tratamento de Dados para Monitoramento e Diagnóstico de Equipamentos Elétricos pelo Método das Descargas Parciais**. Dissertação de Mestrado, Universidade Federal de Minas Gerais, 2002.
- NERI, M.G.G.; GOUVEIA, F.A.O.; COSTA, E.G.; FERREIRA, T.V.; SILVA, K.M.: **Utilização de Redes Neurais na Identificação de Descargas Parciais**. II Simpósio Brasileiro de Sistemas Elétricos, Belo Horizonte, Abril 2008.
- SILVA, G.C.: **Descargas Parciais Estimuladas por Raios-X Contínuo e Pulsado em Materiais Dielétricos: Similaridades e Diferenças**. Tese de Doutorado, Universidade Federal do Paraná. Curitiba, 2005.
- SILVA, K.M.: **Aplicação de Redes Neurais Artificiais na Análise de Ocorrência em Sistemas Elétricos de Potência**. Trabalho de Conclusão de Curso, Universidade Federal de Campina Grande, Fevereiro 2004.
- SIMON, H.: **Redes Neurais: Princípios e Práticas**. Bookman 2º edição. Porto Alegre, 2001.
- KREUGER, F. H.; GULSKI, E.; KRIVDA, A.: **Classification of Partial Discharges**, **IEEE Transactions on Electrical Insulation**, volume. 28, Dezembro. 1993, páginas. 917-931.