Alex Sandro Romeu de Souza Poletto

FILOSOFIA
ANEXO PROC.505/98

Processos de Auto-Organização nos Modelos de Redes Neurais Artificiais

Marília
2002

ALEX SANDRO ROMEU DE SOUZA POLETTO

PROCESSOS DE AUTO-ORGANIZAÇÃO NOS MODELOS DE REDES NEURAIS ARTIFICIAIS

ALEX SANDRO ROMEU DE SOUZA POLETTO

PROCESSOS DE AUTO-ORGANIZAÇÃO NOS MODELOS DE REDES NEURAIS ARTIFICIAIS

Dissertação apresentada ao Programa de Pós-Graduação em Filosofia, na Área de Concentração em "Ciência Cognitiva e Filosofia da Mente", da Universidade Estadual Paulista - Câmpus de Marília.

Orientadora: Dr.ª Maria Cláudia Cabrini Grácio.

1410001417

Marília

2002

F.F.C. de Filosofia e Ciências
BIBLIOTECA
Campus de Marília

ALEX SANDRO ROMEU DE SOUZA POLETTO

PROCESSOS DE AUTO-ORGANIZAÇÃO NOS MODELOS DE REDES NEURAIS ARTIFICIAIS

COMISSÃO JULGADORA

DISSERTAÇÃO PARA OBTENÇÃO DO GRAU DE MESTRE

Presidente e Orientadora Dr.ª Maria Cláudia Cabrini Grácio.

2º Examinador Dr.º Willem F. Haselager

3º Suplente Dr.ª Maria Eunice Quilici Gonzales

Marília, 20 de janeiro de 2002.

unesp

AGRADECIMENTOS

Esta pesquisa foi realizada com muito esforço, dedicação, carinho e vontade. Afinal, foram quatro anos de muitas aulas, pesquisas, leituras, encontros, seminários e palestras para que fosse possível concluí-la. Sinto-me realizado com o término dessa pesquisa, pois aprendi muito durante a sua execução.

Na preparação dessa dissertação, obtive ajuda de vários professores, amigos, colegas e familiares. Em especial, agradeço a minha esposa Flávia, por ter compreendido o quanto seria importante a realização dessa tarefa; ao Dr. Cosme Damião Bastos Massi, pela ajuda na escolha do tema; à minha orientadora Dr.ª Maria Cláudia Cabrini Grácio, pela grande dedicação e atenção, pois foi nesse momento que comecei a acreditar que iria concluir minha dissertação; à Dr.ª Maria Eunice Quilici Gonzales, a Nice, por ter acreditado em meu potencial; ao Prof.º Ms. Luis Carlos Begosso; ao Prof.º Ms. José Augusto Fabri e sua esposa Gabriela, pelos auxílios prestados; ao Prof.º Ms. Luis Ricardo Begosso, pela compreensão em me liberar de minhas atividades para que eu pudesse me dedicar às pesquisas em tempo integral, não poderia deixar de agradecer à Fundação Educacional do Município de Assis – FEMA, pelo apoio profissional e financeiro.

Gostaria, ainda, de agradecer aos amigos e colegas do curso, pelos conhecimentos partilhados em aulas, palestras, seminários, encontros, discussões e reuniões, pois ajudaram muito na realização desta dissertação.

unesp

Para finalizar, agradeço à Unesp e a seus funcionários, por me proporcionarem a realização desse sonho, o de um dia me tornar um mestre. Enfim, obrigado a todos que, de alguma forma, se empenharam em me ajudar, participando de um grande acontecimento em minha vida.

SUMÁRIO

LISTA DE FIGURAS V
LISTA DE TABELAS VII
INTRODUÇÃO 01
1. Teoria de Auto-organização 07
1.1. O que são Sistemas ? 08
1.2. Fundamentos da Teoria da Auto-organização 10
1.2.1. Sobre o conceito de "Organização" 11
1.2.2. Sobre o conceito de "Auto". 15
1.2.3. Sobre o conceito de processos de Auto-organização 15
1.2.3.1. Condição de Partida 17
1.2.3.2. Elementos Distintos 18
1.2.3.3. Encontro e Interação 19
1.2.3.4. Sujeito Onipotente 19
1.2.3.5. Forma Organizada 20
1.2.4. Exemplos de processos Auto-organizados 20
1.3. Auto-organização Primária 23
1.4. Auto-organização Secundária 24
1.5. Fundamentos de Sistemas Auto-organizados 25
1.6. O Papel da Informação e do Ruído nos Sistemas Auto-organizados 28
1.7. Graus da Auto-organização 30
2. Redes Neurais Artificiais 34
2.1. Fundamentos das Redes Neurais Artificiais 35
2.2. O que motivou as pesquisas sobre as Redes Neurais Artificiais 36
2.3. Unidade de Processamento (Neurônio Artificial) 38
2.4. Conceitos básicos sobre Formação de Padrões Informacionais 51
2.5. Principais Arquiteturas das Redes Neurais Artificiais 53
3. Aprendizagem 58
3.1. Fundamentos de Aprendizagem 58
3.2. Aprendizagem nas Redes Neurais Artificiais 61
3.2.1. Método de Aprendizagem Supervisionada 65
3.2.2. Método de Aprendizagem sem um Supervisor ou por Auto-organização 69
4. Modelos de Redes Neurais Artificiais 73
4.1. Modelos com Aprendizagem Supervisionada 73
4.1.1. *Perceptron de Camada Única* 73
4.1.2. Algoritmo *Backpropagation* e *Perceptron de Múltiplas Camadas* 77
4.2. Modelos com Aprendizagem Não Supervisionada 80
4.2.1. *SOFM* de Kohonen 81
4.2.2. Mecanismo da rede *SOFM* 84
4.2.3. Algoritmo da rede *SOFM* 85
4.3. Treinamento de uma Rede Neural Artificial *Backpropagation* para reconhecimento e classificação de padrões informacionais 89
4.3.1. Descrição dos dados 90
4.3.2. Definição das camadas de entrada, intermediária e de saída 93

4.3.3. Definição das Arquiteturas 93
4.3.4. Configurações das Arquiteturas 95
4.3.5. Análise dos Resultados 99
4.4. Treinamento de uma Rede Neural Artificial *SOFM* de Kohonen para reconhecimento e classificação de padrões informacionais 102
4.4.1. Configuração da rede para as simulações 103
4.4.2. O Treinamento 104
5. Análise do Modelo *Perceptron de Múltiplas Camadas* e do Modelo *SOFM* de Kohonen em relação à Processos da Auto-organização 110
5.1. O modelo *Perceptron de Múltiplas Camadas – Backpropagation* (aprendizagem supervisionada) 112
5.2. O modelo *SOFM* de Kohonen (aprendizagem não supervisionada) 115
CONSIDERAÇÕES FINAIS 122
PROJEÇÕES FUTURAS 125
BIBLIOGRAFIA 127

LISTA DE FIGURAS

Figura 1. Exemplo de um sistema único, composto por partes integrantes e interdependentes (Filho e D'Ottaviano, 2000, pp. 283-306)..... 09
Figura 2. Esquema de um processo de organização, um sistema (um TODO), composto de sub-sistemas (várias PARTES)..... 12
Figura 3. Exemplo da existência de condicionalidade (Debrun, 2000, p.85)14
Figura 4. Função Limite Ríspido (*Hard Limiter*) (Haykin, 2001, p.39 e Lippmann, 1987, p. 14) 42
Figura 5. Função Limite Lógico (*Threshold Logic*) (Haykin, 2001, p.39 e Lippmann, 1987, p. 14) 43
Figura 6. Função Sigmóide (*Sigmoid*) (Haykin, 2001, p.39 e Lippmann, 1987, p. 14)..... 43
Figura 7. Composição de uma unidade de processamento (Braga, 1998, p. 8) 44
Figura 8. Composição detalhada das interconexões entre várias unidades de processamento..... 46
Figura 9. Exemplo de uma rede neural artificial de 4 camadas, sendo 1 camada de entrada, 2 intermediárias e 1 de saída (Kovács, 1996, p. 40)..... 47
Figura 10. Exemplo de uma unidade de processamento ou uma *unidade neurônio-símile* (UNS) 50
Figura 11(a). Modelo de Simulação com sinais fornecidos pelo supervisor..... 52
Figura 11(b). Modelo de Simulação com sinais provindos do meio ambiente... 52
Figura 12(a). Camada Única – Acíclica - Totalmente Conectada (Braga, 1998, p. 11)..... 55
Figura 12(b). Múltiplas Camadas – Acíclica - Parcialmente Conectada (Braga, 1998, p. 11)..... 55
Figura 12(c). Múltiplas Camadas – Cíclica - Parcialmente Conectada (Braga, 1998, p. 11)..... 56
Figura 12(d). Camada Única – Cíclica - Totalmente Conectada (Braga, 1998, p. 11)..... 56
Figura 13. Exemplo de uma rede neural artificial com método de aprendizagem supervisionada (Haykin, 2001, p. 88)..... 67
Figura 14. Esquema de uma rede neural artificial com método de aprendizagem por reforço (Haykin, 2001, p. 90)..... 70
Figura 15. Esquema de uma rede neural artificial com método de aprendizagem não supervisionada (Haykin, 2001, p. 91) 71
Figura 16. Modelo de uma rede neural artificial *Perceptron de Camada Única* (Braga, 1998, p. 28)..... 75
Figura 17. Esquema de uma rede neural artificial *Perceptron de Múltiplas Camadas* 79
Figura 18. Modelo de uma rede neural artificial *Perceptron de Múltiplas Camadas* 80
Figura 19. Tipos de estímulos produzidos de acordo com a função chapéu mexicano (Braga, 1998, p. 105)..... 83

Figura 20. Exemplo de uma rede *SOFM* de Kohonen típica (bidimensional) (Braga, 1998, p. 106) 87
Figura 21. Topologia de vizinhança de acordo com o tempo variável (Braga, 1998, p. 110) 88
Figura 22. Exemplo de uma rede de aprendizado competitivo (Gonzales, 1996, pp. 279-282) 89
Figura 23. Configuração da rede para o primeiro treinamento - *Backpropagation* 96
Figura 24. Configuração da rede para o segundo treinamento - *Backpropagation* 97
Figura 25. Configuração da rede para o terceiro treinamento *Backpropagation* 98
Figura 26. Configuração inicial da rede para o treinamento, extraída do software de Simulação da Rede Neural de Kohonen 105
Figura 27. Configuração da rede após 1000 iterações, extraída do software de Simulação da Rede Neural de Kohonen 106
Figura 28. Configuração da rede após 2000 iterações, extraída do software de Simulação da Rede Neural de Kohonen 106
Figura 29. Configuração da rede após 3000 iterações, extraída do software de Simulação da Rede Neural de Kohonen 107

LISTA DE TABELAS

Tabela 1. Atributos utilizados nos treinamentos - *Backpropagation* 92
Tabela 2. Arquiteturas utilizadas nos treinamentos - *Backpropagation* 93
Tabela 3. Resultados do primeiro treinamento - *Backpropagation* 96
Tabela 4. Resultados do segundo treinamento - *Backpropagation* 98
Tabela 5. Resultados do terceiro treinamento - *Backpropagation* 99
Tabela 6. Resultados dos treinamentos - taxa de erro 0.05 - *Backpropagation* .100
Tabela 7. Resultados dos treinamentos - taxa de erro 0.10 - *Backpropagation* .100

INTRODUÇÃO

As redes neurais artificiais surgiram na década de 40, a partir da idéia de McCulloch e Pitts de simular ou, até mesmo, incorporar características do cérebro humano em máquinas e computadores. A idéia era criar um neurônio artificial, que simulasse um neurônio biológico, incorporando naquele algumas características e propriedades pertinentes a este último. A esse neurônio artificial foi dado o nome de *unidade de processamento*.

No final da década de 40, mais precisamente em 1949, o psicólogo Donald Hebb fez importantes contribuições às redes neurais artificiais, ao elaborar uma teoria baseada no processo de aprendizagem que ocorre no cérebro humano. Esta teoria serviu de base para a aprendizagem das redes neurais artificiais.

Posteriormente, nos anos 50, surgiu um dos mais significantes resultados apresentados sobre as redes neurais artificiais, o modelo *Perceptron*, criado e desenvolvido por Rosenblat em 1957. Esse modelo possuía um procedimento simples de treinamento que convergia para uma resposta, se a solução para o problema proposto de fato existisse, isto é, caso o problema fosse linearmente tratável, em caso contrário, não convergia.

Minsky e Papert (1969) na década de 60, descobriram algumas falhas no modelo *Perceptron*, observando e alertando sobre algumas tarefas que o mesmo não era capaz de executar: esse modelo somente resolvia problemas linearmente separáveis, nos quais uma linha reta é usada para separar os estímulos

de entradas em apenas duas classes. Minsky e Papert defenderam também a hipótese de que o modelo *Perceptron* somente convergia nos modelos de *camada única* de unidades de processamento. Modelos conhecidos como de *múltiplas camadas* não apresentavam, necessariamente, convergência para soluções, mesmo quando elas existiam.

Em virtude das declarações feitas por Minsky e Papert, na década de 60, criticando o modelo *Perceptron*, as pesquisas sobre redes neurais artificiais ficaram estagnadas na década de 70.

Na década de 80, as pesquisas sobre redes neurais artificiais renasceram, especialmente após a publicação do artigo de John Hopfield (1982), que mostrava a relação entre modelos de memória associativa e sistemas físicos, tornando-se atualmente linha de pesquisa em várias universidades.

Os modelos de redes neurais artificiais podem hoje ser divididos, quanto ao método básico de aprendizagem, em dois tipos: redes neurais artificiais que utilizam o método de *aprendizagem supervisionada*; e redes neurais artificiais que utilizam o método de *aprendizagem não supervisionada.*

Teoricamente, encontram-se na literatura inúmeras discussões sobre o uso dos conceitos básicos da *teoria da auto-organização* e de *sistemas auto-organizados* em modelos de redes neurais artificiais. Pretende-se, desse modo, averiguar se de fato tais modelos apresentam processos de auto-organização, e investigar em que medida os modelos de redes neurais artificiais são capazes de modelar e de simular esses processos.

Segundo Debrun (1996), o termo *auto-organização* tem sido caracterizado de maneiras diferentes, por trás das quais existe uma idéia comum, uma intuição instigante de que novas estruturas possam emergir mais ou menos espontaneamente em certos domínios. Baseando-se nessa caracterização da auto-organização, esta investigação se iniciará, tendo como foco de análise os modelos de redes neurais artificiais.

Os processos de auto-organização dividem-se em dois tipos: a *auto-organização primária* e a *auto-organização secundária*. Supõe-se que esses dois tipos de auto-organização são utilizados por vários modelos de redes neurais artificiais. Resta esclarecer até que ponto esses processos estão incorporados, ou seja, de fato implantados nos modelos de redes neurais artificiais.

O interesse em se saber ***se processos de auto-organização ocorrem nos modelos de redes neurais artificiais*** surgiu, principalmente, em razão da preocupação em se esclarecer em que medida os fundamentos teóricos dos processos de auto-organização se aplicam aos *modelos de redes neurais artificiais*.

Neste contexto, o objetivo da presente pesquisa é investigar os processos da auto-organização *primária* e *secundária* em relação a alguns modelos de redes neurais artificiais. Esta investigação mostrará se processos de auto-organização, por meio de seus dois tipos, ocorrem nos processos dos modelos de redes neurais artificiais, focalizando o modelo *SOFM* de *Kohonen*, que utiliza o método de aprendizagem não supervisionada. Antes, será feito, também, um estudo de dois modelos que se servem do método de aprendizagem

supervisionada. Trata-se do modelo *Perceptron de Camada Única* e do modelo *Perceptron de Múltiplas Camadas* (*Backpropagation*).

Um outro objetivo desta pesquisa é aprimorar o conhecimento em relação aos modelos de redes neurais artificiais que se utilizam dos métodos de aprendizagem supervisionada e não supervisionada. Para tal investigação, torna-se primordial o estudo de alguns assuntos básicos que estão diretamente ligados ao objetivo global desse trabalho. Portanto, serão introduzidos os conceitos de *Auto-organização*, *Sistemas Auto-organizados*, *Redes Neurais Artificiais*, *Métodos de Aprendizagem*, *Modelos de Redes Perceptron de Camada Única* e de *Múltiplas Camadas (Backpropagation)* e do *Modelo SOFM* de Kohonen (*Mapas de Características Auto-organizantes)*.

No capítulo 1, serão apresentados conceitos básicos sobre *Sistemas*, fundamentos teóricos pertinentes à *Auto-organização*, passando-se obrigatoriamente pelas noções básicas do termo "Organização" e do prefixo "Auto", e pelos conceitos básicos de *processos de auto-organização*, bem como pelas noções de *Condição de Partida*, *Elementos Distintos*, *Encontro e Interação*, *Sujeito Onipotente* e *Forma Organizada*, já que são estas as noções centrais subjacentes aos processos de auto-organização. Serão vistos, também, os dois tipos de auto-organização, a *primária* e a *secundária*, bem como os fundamentos teóricos associados à noção de *Sistemas Auto-organizados*. Serão apresentados o *papel da informação* e do *ruído* em sistemas auto-organizados e, também, os graus nos quais os processos da auto-organização podem ou não ocorrer.

A *motivação* da continuidade dos estudos sobre *Redes Neurais Artificiais*, bem como alguns dos seus fundamentos teóricos serão os tópicos ressaltados no capítulo 2, passando-se pelo conceito de *neurônio artificial* e pelos conceitos básicos de *formação de padrões informacionais*. Consideraremos ali também, como funciona uma rede neural artificial, bem como as *principais arquiteturas* dos modelos de redes neurais artificiais.

Os *Processos de Aprendizagem* serão o tema do capítulo 3, iniciando-se a investigação com os fundamentos teóricos de *aprendizagem*, passando-se pelos dois métodos de aprendizagem que podem ocorrer nos modelos de redes neurais artificiais: a *aprendizagem supervisionada* e a *aprendizagem não supervisionada*. As características dos modelos de redes neurais artificiais supervisionados e dos modelos de redes neurais não supervisionados ou auto-organizados serão assunto aqui também.

No capítulo seguinte, serão estudados dois modelos de redes neurais artificiais que utilizam processos de aprendizagem supervisionada e um modelo que utiliza processos de aprendizagem não supervisionada. Como exemplo de modelos de redes neurais artificiais com aprendizagem supervisionada, serão apresentados os modelos *perceptron de camada única* e o *perceptron de múltiplas camadas*, este último conhecido por *algoritmo de retropropagação* (*backpropagation*). Como modelos de redes neurais artificiais com aprendizagem não supervisionada ou por auto-organização, será mostrado o modelo *SOFM* de Kohonen – *Mapas de Características Auto-organizantes*. Serão realizados ainda dois treinamentos para reconhecimento e classificação de

padrões. No primeiro treinamento, será utilizado o modelo *perceptron de múltiplas camadas*, com aprendizagem supervisionada, e, no segundo, o modelo *SOFM* de Kohonen, com aprendizagem não supervisionada.

O quinto e último capítulo traz uma análise dos dois modelos: com *aprendizagem supervisionada* e com *aprendizagem não supervisionada*, anteriormente analisados. É neste capítulo que se examinará se *processos auto-organizados* ocorrem nos modelos de redes neurais artificiais, e em caso afirmativo, até que ponto isso acontece.

As considerações finais não terão caráter definitivo e consistirão de uma síntese dos resultados obtidos e também algumas projeções para o futuro.

1

Teoria da Auto-organização

Este capítulo apresenta uma parte substancial do embasamento teórico necessário para se atingir o objetivo central desta pesquisa, uma vez que se pretende observar em que medida os modelos de redes neurais artificiais são capazes de modelar e simular processos de auto-organização.

Para a realização desta proposta, dividiu-se o capítulo em seis seções. Na primeira seção será apresentado o que vem a ser *Sistemas*.

Na segunda seção, serão enfocados os fundamentos da teoria da *auto-organização*, precedidos da explicação das noções básicas do termo *organização* e do prefixo *auto*, bem como os conceitos básicos dos processos auto-organizados. Igualmente, terão destaque as noções centrais de que a auto-organização faz uso: *condição de partida*, *encontro e interação*, *elementos distintos*, *sujeito onipotente* e *forma organizada*.

Na seção seguinte será apresentado o primeiro tipo de auto-organização, a *auto-organização primária*.

O segundo tipo de auto-organização, a *auto-organização secundária*, será abordado na quarta seção.

Na quinta seção serão enfocados fundamentos de *sistemas auto-organizados*, fazendo uso da primeira seção que trata dos conceitos básicos de *sistemas*, visto que os processos de auto-organização primária constituem um sistema e os processos de auto-organização secundária, por meio da interação de

elementos, fazem uso desse sistema primariamente constituído, a fim de instaurar sistemas mais evoluídos, *sistemas que se auto-organizam*.

Na seção seis, discutir-se-á de que maneira a *informação* influencia nos processos dos sistemas auto-organizados e qual é a função do *ruído informacional*, por meio de suas duas espécies: de *catalisador* e de *afrouxador de canais de comunicação*, durante os processos nos sistemas auto-organizados.

Na última seção, serão apresentados os *graus da auto-organização* que podem ou não ocorrer em processos de auto-organização.

1.1. O que são Sistemas ?

De modo geral, pode-se entender um sistema como uma entidade unitária constituída por um grupo de elementos ou sub-entidades, que se mantêm em constante interação (comunicação), garantindo sua própria existência, sobrevivência, identidade.

Segundo Grisson (1999, p. 13), sistema é "um conjunto de elementos em interação, ou ainda, um conjunto de partes integrantes e interdependentes que, em conjunto, formam um todo unitário com determinado objetivo e efetuam determinada função".

A Figura 1, de própria autoria, ilustra a formação de um sistema único, isto é, de um todo composto por partes integrantes e interdependentes. Analisando esta figura, observa-se que as divisões do circulo interno constituem as partes (componentes) separáveis do sistema que, ao interagirem, criam partes dependentes, as quais acabam por formar o sistema propriamente dito. As setas

cm 1 2 3 4 5 6 7 unesp 1 12 13 14 15 16 17 18

bidirecionais de cor azul, constituem as interfaces ou canais de comunicação (interação) entre as partes, resultando na formação do todo. As setas bidirecionais de cor marrom, realizam a comunicação do sistema (todo) com o meio ambiente externo, caracterizado na figura pelo círculo externo.

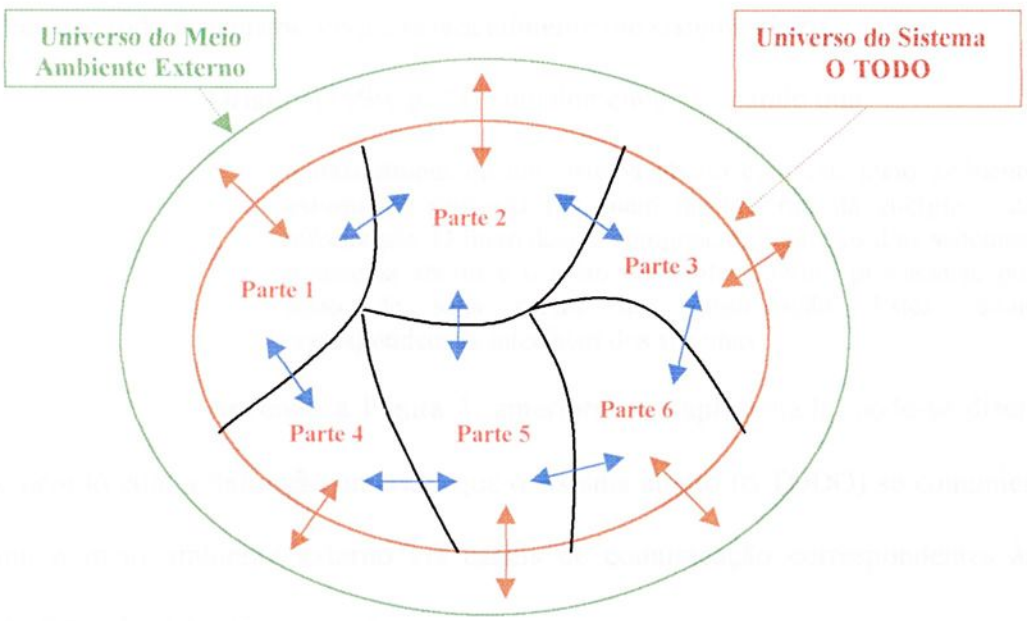

Figura 1. Exemplo de um sistema único, composto por partes integrantes e interdependentes inspirada em Filho e D'Ottaviano (2000, pp. 283-306).

Conforme Grisson (1999), pode-se classificar os sistemas em dois tipos: *fechados* e *abertos*. Os sistemas fechados "são assim chamados por serem considerados isolados de seu ambiente" (Grisson, 1999, p. 14). Assim sendo, em qualquer sistema fechado, o estado final é inequivocadamente determinado por suas próprias condições iniciais e internas.

Quanto aos sistemas abertos, Grisson (1999, p. 14) os define como aqueles que "mantêm um contínuo fluxo de entrada e saída de informações,

que se conserva mediante a construção e a decomposição de componentes, nunca estando, enquanto vivo, em um estado de equilíbrio químico e termodinâmico"

Assim sendo, em qualquer sistema aberto, o estado final pode ser alcançado, partindo-se de diferentes condições iniciais e diversas maneiras. O empenho na organização do sistema visa sua sobrevivência. Portanto, pode-se dizer que todo organismo vivo é essencialmente um sistema aberto.

Grisson (1999, p. 14) complementa, afirmando que

> o intercâmbio de um sistema aberto com seu meio ambiente externo se processa por meio da matéria, da energia e da informação. O fluxo desses componentes entre os dois sistemas, o sistema aberto e o meio ambiente externo, processa-se por meio de seus canais de comunicação. Estes canais correspondem às interfaces dos sistemas.

Revendo a Figura 1, anteriormente apresentada, pode-se dizer, de acordo com a definição anterior, que o sistema aberto (o TODO) se comunica com o meio ambiente externo via canais de comunicação correspondentes às interfaces do sistema.

1.2. Fundamentos da Teoria da Auto-organização

Um dos objetivos deste trabalho é averiguar se processos auto-organizadores ocorrem em alguns modelos de redes neurais artificiais. Esta investigação apresenta, de início, os fundamentos e os conceitos encontrados na literatura acerca da idéia de auto-organização, que representa o fio condutor no desenvolvimento desta pesquisa.

Tratando-se da teoria de auto-organização é oportuno apresentar os conceitos de *sistema*, de *organização* e de *auto*, uma vez que tanto o prefixo *auto* quanto o termo *organização* possuem mais de um significado.

1.2.1. Sobre o conceito de "Organização"

Considerando que neste trabalho trata-se basicamente de processos artificiais entre elementos, descarta-se desde já a possibilidade de que a noção de *organização*, aqui referida, tenha alguma ligação com aquela usada na definição de empresa, fundação, instituição ou indústria.

Segundo Debrun (1996, *XXXIV*), a noção de *organização* aplica-se aos casos nos quais ocorrem "ajustes e aprimoramentos entre elementos distintos que componham processos artificiais e/ou biológicos". Esses elementos podem ser os membros de grupos de várias situações, como por exemplo: grupos sociais, grupos políticos, grupos acadêmicos, grupos esportivos, etc. Nesses casos de processos de organização, podem-se citar reuniões de grupos políticos para a discussão de assuntos que dizem respeito às leis governamentais ou normas parlamentares, reuniões ou encontros de grupos acadêmicos para estudos de assuntos científicos, etc. Quando os elementos de um grupo se encontram, podem ocorrer ajustes e aprimoramentos entre eles. Por conseguinte, as reuniões, os encontros, propiciam a organização do grupo de elementos, que se torna, assim, um sistema organizado. Existe, porém, a necessidade de que esses elementos, que são as partes, se comuniquem entre si para compor o todo, o qual, por sua vez, poderá vir a ser uma parte de um outro todo maior e mais complexo.

A figura a seguir, de própria autoria, ilustra genericamente um esquema de sistema (todo) e sub-sistemas (partes):

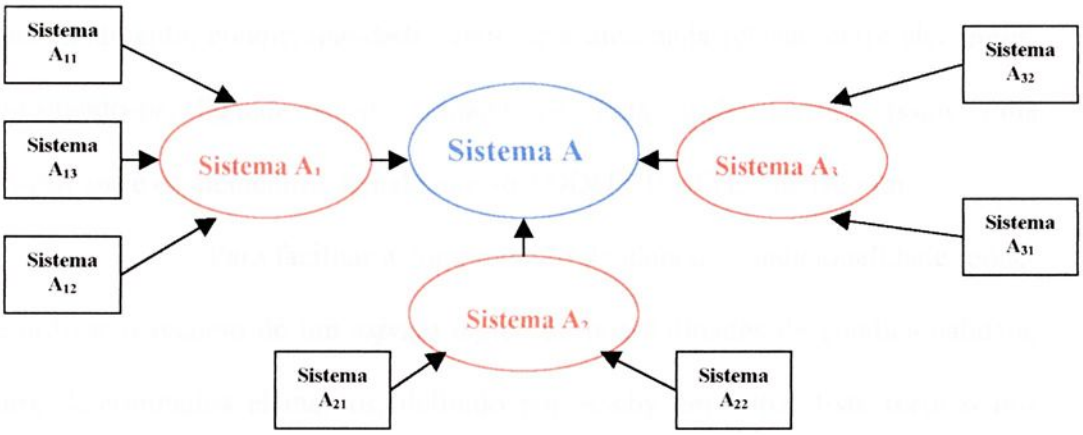

Figura 2: Esquema de um processo de organização, um sistema (um TODO), composto de sub-sistemas (várias PARTES).

Analisando-se a Figura 2, verifica-se que o Sistema A possui três sub-sistemas definidos como A_1, A_2 e A_3, que, por sua vez, possuem vários outros sub-sistemas. O sistema A_1 possui três sub-sistemas: A_{11}, A_{12} e A_{13}; o sistema A_2 possui dois sub-sistemas: A_{21} e A_{22}; e o sistema A_3 também possui dois sub-sistemas: A_{31} e A_{32}. Na visão geral, o sistema A, pode ser definido como o TODO mais complexo, maior, visto que a ele estão ligados os sistemas A1, A2 e A3, definidos aqui como PARTES desse TODO mais complexo. Mas essas PARTES acabam se tornando o TODO mais simples, visto que esses sistemas também possuem sistemas que estão diretamente ligados a eles. Portanto, esses sistemas (A1, A2 e A3) realizam os dois papéis, de PARTE e de TODO, mas um TODO mais simples, visto que suas PARTES não possuem subsistemas.

Ainda sobre organização, mas segundo a concepção de Ashby (1962), para que realmente se compreenda o sentido mais específico de organização, é primordial que se esclareça a noção de *condicionalidade*. Resumidamente, condicionalidade refere-se a uma dada relação entre elementos, conectando-os, ligando-os, direcionando-os. Esta noção possui, assim, uma relação entre os elementos, semelhante ao TODO e PARTES de Debrun.

Para facilitar a compreensão da idéia de condicionalidade, pode-se utilizar o recurso de um *espaço lógico* de possibilidades de condicionalidade entre determinados elementos, definido por Ashby em 1962. Este recurso nos ajudará a averiguar quando há ocorrência da condicionalidade entre elementos. No exemplo dado por Ashby (1962), os elementos A e B se encontram num sistema *X*. Suponhamos que no momento m_1 temos o elemento A em um certo estado a_1 e o conjunto de todos os estados possíveis do elemento B dado por $\{b_1, b_2, b_3, ..., b_n\}$. Por outro lado, se no momento m_2 há uma probabilidade igual de ocorrência para todos os estados possíveis de B, então, conforme Ashby, pode-se dizer que não está havendo nenhuma comunicação no sentido de A condicional a B, ou seja, *não há restrição* realizada por A no conjunto de estados em que B possa ser encontrado, isto é, 'não há condicionalidade' nesta concepção. Isto é, se não existe uma comunicação entre dois sistemas, não existe também a possibilidade de que um sistema possa condicionalizar ou restringir uma situação, um estado, uma ação, etc. Imaginando-se um segundo sistema *Y*, em que o momento m_1 seja idêntico ao momento m_1 do sistema *X*, conforme Ashby, pode-se ter, por exemplo, no momento m_2 de *Y*, a possibilidade de ocorrência de apenas b_2 ou b_3. Neste

caso, existe condicionalidade, ou seja, o elemento A *restringiu* as possibilidades de ocorrência do elemento B.

A figura abaixo ilustra a idéia de condicionalidade, segundo Ashby (1962), utilizando o recurso descrito anteriormente.

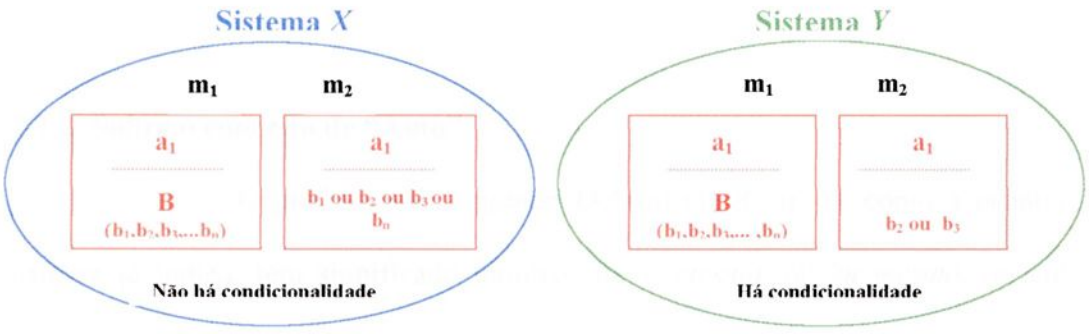

Figura 3: Exemplo da existência de condicionalidade, extraída de Debrun (2000, p. 85).

Ashby acredita que é essencial que se entenda organização como um todo composto de partes (ver Figura 2), ou seja, que a partir de sub-sistemas (partes) possam-se formar sistemas mais complexos (o todo).

De outra parte, é viável discutir-se a utilidade da condicionalidade, como por exemplo, se uma organização é *boa*, segundo certos aspectos, ou *ruim* segundo outros tantos.

Segundo Ashby (1962, p. 262), "uma organização é considerada boa se consegue manter suas variáveis essenciais por meio de um certo dinamismo no tempo e no espaço, e ruim se não consegue isto". Em outras palavras, uma boa organização é aquela que promove a estabilidade (sobrevivência) do sistema. Se houver total dispersão, redução ou separabilidade, não há como se falar em organização.

Portanto, conforme Ashby (1962, p. 266), não parece haver boa organização que seja válida para todos os casos, em um sentido absoluto. Uma boa organização, num determinado contexto, pode ser ruim se transferida para uma outra situação, um outro contexto; o que eqüivale a dizer que a boa organização o é sempre para uma determinada situação particular.

1.2.2. Sobre o conceito de "Auto"

O prefixo *auto*, segundo Debrun (1996, p. 6), como a própria palavra já indica, tem significado intuitivo de *si mesmo*, ou *eu mesmo*, ou *ele mesmo*, ou *nós mesmos*.

Segundo Debrun (1996, p. 25) o prefixo "Auto" passa a idéia de que o sujeito auto-organizador permanece "nele próprio", ou seja, de que ele próprio efetua um trabalho de *si para si*. Tanto a *condição de partida*, como o *objeto de aplicação*, bem como os *mecanismos e recursos* utilizados situam-se dentro de um mesmo sistema. Assim, a referência a qualquer propriedade precedida pelo prefixo *Auto* indica que ele se aplica ao próprio sistema em estudo.

1.2.3. Sobre o conceito de processos de Auto-organização

Após a conceituação do termo *Organização* e do prefixo *Auto*, pode-se, então, apresentar os conceitos básicos relativos aos *processos de auto-organização.*

Conforme Debrun (1996, parte I), qualquer sistema que possa entrar em funcionamento, manter-se em funcionamento e organizar-se sem

unesp

depender de ajuda externa, isto é, sem depender de elementos que não estejam dentro do contexto do sistema, mesmo estando aberto ao meio ambiente externo é denominado um *sistema auto-organizado* ou um *sistema* que apresenta aspectos da *auto-organização*, já que ele realizou um processo *de si para si*.

O fato de o sistema não depender necessariamente de ajuda externa, não o torna um sistema fechado ao meio ambiente externo. O que se focaliza aqui é que o intercâmbio com o meio ambiente externo deverá ter um grau de incidência menor do que o grau de incidência do próprio sistema em questão.

Segundo Debrun (1996, p.21), um processo de auto-organização é aquele que

> se inicia por meio do rompimento com uma continuidade e que é promovido pelo encontro e interação entre elementos distintos ou semi-distintos, sem a presença de um sujeito onipotente ou de determinações preexistentes, e que culmina com a cristalização de uma forma final organizada.

A análise das definições de Debrun indicam que, para a auto-organização ser possível, esses processos devem conter mecanismos que possibilitem seu próprio funcionamento, manutenção e organização, sem depender unicamente das informações que eles recebem de sistemas, processos ou elementos externos; a saber, tudo que ocorre fora do seu ambiente interno, de seu domínio, não pode causar mudanças, nem pode influenciar os elementos envolvidos nesse processo durante suas tomadas de decisões.

De maneira geral, a partir das definições de Debrun, pode-se dizer que esses sistemas se tornam processos auto-organizados, ou seja, processos

que têm seu próprio início e seu próprio término, organizando-se a si mesmos, já que dependem única e exclusivamente deles mesmos para se organizarem.

O fato de esses processos dependerem única e exclusivamente deles mesmos para se organizarem, não quer dizer que não sejam sistemas abertos ao meio ambiente externo. Entretanto, a organização de modo geral deverá ser realizada pelo próprio sistema em questão e não por influência exclusiva do meio ambiente externo.

Algumas noções são primordiais para a ocorrência de um processo de auto-organização, a saber: *Condição de Partida*, *Encontro e Interação*, *Elementos Distintos*, *Sujeito Onipotente* e *Forma Organizada* (Debrun, 1996).

Nas subseções seguintes serão tratadas individualmente cada uma das noções subjacentes à ocorrência de processo de auto-organização.

1.2.3.1. Condição de Partida

Debrun (1996) afirma que, para que se tenha início um processo de auto-organização, há a necessidade de uma *condição de partida* que quase sempre se origina da quebra ou rompimento de um processo original contínuo. Após tal rompimento é que surge um novo processo que dá início à auto-organização.

Desta forma, ressalta Debrun que a auto-organização possui seu próprio começo. O processo obtido pela quebra ou rompimento de um processo original contínuo de uma condição de partida, será considerado seu próprio

começo. Isto significa que houve um desvinculamento ou perda de todo o passado e um ganho de liberdade em relação ao passado quando da ligação entre os elementos distintos que estarão envolvidos nesse processo de auto-organização.

1.2.3.2. Elementos Distintos

Elementos distintos são os integrantes principais dos processos de auto-organização, visto que sem eles, não existiriam tais processos. Os grupos são compostos por elementos distintos, nos quais, por meio dos encontros e das interações, dar-se-ão os processos de auto-organização.

Em processos de auto-organização faz-se necessário que os elementos sejam realmente distintos, isto é, que não exista entre eles qualquer vínculo preexistente de dependência, garantindo a existência do acaso na condição de partida.

No início do processo, como os elementos não se conhecem, prevalece o papel do acaso. Durante os demais encontros, com a crescente interação, maior vínculo e entrosamento entre os elementos, o papel do acaso acaba por diminuir progressivamente, prevalecendo, então, a interação.

1.2.3.3. Encontro e Interação

Os encontros ocorridos, por obra do acaso, entre os elementos distintos levam-nos a ficar mais próximos uns dos outros, sem, entretanto, uma ordem definida. Todavia, devido ao meio de ajustes de posicionamentos, os elementos podem acabar por alcançar alguma ordem. Nos momentos em que

ocorrem os *encontros* começam a surgir as *interações* entre eles. Quanto mais interações ocorram entre esses elementos, menor será o papel do acaso. Isto se deve a partir do momento em que os elementos passam a realizar os encontros e a interagir, o acaso deixa de existir. Entretanto, Debrun lembra que o encontro inicial deve ocorrer por obra do acaso.

Portanto, encontro e interação são condições fundamentais para a existência dos processos de auto-organização.

1.2.3.4. Sujeito Onipotente

Para que um processo artificial entre elementos distintos se torne um processo auto-organizado, não pode existir um *sujeito onipotente* ou um *supervisor*. Um sujeito onipotente é um elemento que, em um processo biológico ou artificial, domina ou influencia outros elementos. Desse modo, nesses processos, todos os elementos envolvidos deverão estar no mesmo grau de influência, ou seja, todos têm os mesmos direitos e privilégios perante os demais elementos.

Em processos de auto-organização não existe a presença de um elemento supervisor, isto é, não há nenhum elemento que dê ordens, que seja dominante ou comandante do grupo. Caso isso ocorra, não se trata mais de processos de auto-organização, e sim de *processos de hetero-organização*, ou seja, processos em que há a presença de um ou mais elementos dominantes ou supervisores, que comandam os demais elementos integrantes do grupo.

1.2.3.5. Forma Organizada

A última noção aqui apresentada é aquela referente à cristalização de uma *forma organizada*, isto é, a constituição de um estado relativamente estável que tende a estabilizar-se. Neste estado, consegue-se evidenciar a diferença entre o passado e o presente. A constituição de uma forma organizada é aquela que irá confirmar se o processo de auto-organização já foi ou está para ser alcançado.

Nas noções apresentadas, segundo Debrun (1996, p. 13), encontram-se processos de auto-organização toda vez que,

> um processo de encontro e interação entre elementos distintos tende, sem intervenção de um supervisor, pelo menos um supervisor onipotente, à constituição de uma forma global (ou todo). Sedimenta progressivamente seu próprio atrator e portanto, se cria a si próprio enquanto sistema, resulta essencialmente da própria interação e só em grau menor de sua condição de partida ou do intercâmbio com o ambiente.

1.2.4. Exemplos de processos Auto-organizados

Como exemplo de um processo de Auto-organização, imaginemos que seja criado um grupo para pesquisar um determinado assunto com pessoas totalmente distintas, ou seja, pessoas que não possuem entre elas nenhum tipo de dependência. Segundo os parâmetros da auto-organização, cada uma das noções anteriormente mencionadas teriam as seguintes caracterizações:

- **Condição de Partida**

A formação de um grupo cujos integrantes possuem liberdade uns para com os outros, por meio da quebra ou desconexão com seus respectivos passados.

- **Elementos Distintos**

Os integrantes do grupo, cada um, inicialmente, sem qualquer tipo de dependência em relação aos outros integrantes são os elementos distintos.

- **Encontro e Interação**

Os integrantes do grupo passam a reunir-se em encontros e começam a interagir como grupo; o papel do acaso diminui cada vez mais e a interação passa a dominar totalmente o grupo.

- **Sujeito Onipotente (Pessoa ou elemento Onipotente)**

O referido grupo de pesquisa não tem nenhum integrante dominante ou supervisor, já que se trata de um grupo auto-organizado e não de um grupo hetero-organizado. Os integrantes do grupo têm os mesmos direitos e privilégios perante os demais integrantes. O sujeito é adaptado de acordo com o ambiente em questão; em nosso exemplo, as pessoas (elementos) é que fazem o papel do sujeito.

- **Forma Organizada**

Ao final do processo, tem-se a formação de um grupo auto-organizado, em que cada integrante cumpre o seu papel. O resultado final pode ser um grande projeto ou até mesmo um simples projeto, indicativo de uma evolução mediante o que foi proposto, inicialmente, na criação do grupo de pesquisa.

Como outro exemplo de auto-organização, pode-se citar aquele apresentado por Debrun (1996, Prefácio, p. xxxiv), referente a um jogo de futebol. Nele acontecem, às vezes, inúmeras situações nas quais se estabelece (até o fim do jogo) quer a liderança incontestada de uma das equipes, quer um marasmo

generalizado sem que a relação energética das forças oponentes conduzam a um outro desfecho. Parece haver uma aceitação de todos, mesmo que não haja clara consciência do que está ocorrendo.

Nesta situação não há um ajuste planejado. Para este sujeito coletivo não estão programadas certas ações (movimentos, atitudes, etc.) que os jogadores devam exercer. Sabe-se que um determinado jogador, em um curto espaço de tempo, deverá tomar certas decisões em razão das situações em que seu parceiro de equipe se encontra e também da situação dos jogadores adversários. Existem inúmeras situações tanto com relação aos jogadores (parceiros ou adversários) quanto em relação à situação do juiz, ao tipo de clima ambiente (chuva, sol, noite, calor, altitude, etc.), às ordens do treinador, à vontade dos torcedores. Essas situações podem causar mudanças de certa forma organizatória, no sentido daquilo que cada equipe pretende alcançar, naquele momento do jogo.

Pode-se dizer que o objetivo de cada equipe é ganhar o jogo, mas não se pode prever o resultado. E já que são tantas as situações que se deve levar em consideração, não sendo possível cruzá-las para se chegar a uma resposta definitiva e correta do resultado.

1.3. Auto-organização Primária

O conceito de *auto-organização primária* está relacionado ao "encontro e embate entre os elementos 'realmente distintos' reunidos especialmente por obra do acaso, dentro de um quadro de interação, mas sem supervisor nem objetivo global a alcançar" (Debrun, 1996, p. 20).

Desse modo, conforme Debrun, a auto-organização primária ocorre por meio de encontros entre elementos que não tiveram nenhuma relação no passado, que não possuem relação de dependência. Esse encontro se dá ao acaso, ou seja, sem nada definido, marcado, combinado. Esses elementos interagem sem que haja um elemento dominante, nem objetivo ou forma organizada a alcançar. Ao final do encontro, porém, poderá surgir alguma forma ou objetivo.

Como já mencionado na subseção 1.2.3.4, esses elementos possuem os mesmo privilégios, ou seja, não existe nenhum elemento que comandará o encontro, não há chefe algum, todos estão no mesmo grau de domínio.

O processo de auto-organização primária aplica-se ao exemplo do grupo para a elaboração de um projeto de pesquisa sobre um determinado assunto. Não temos nada formado, a não ser o próprio grupo, mas no final desse trabalho poderá surgir um projeto de pesquisa sobre um determinado assunto. Tal projeto seria determinado por: cronogramas, metodologias, pontos de partida, objetivos, bibliografia, etc., absorvidos pelo grupo e a ele aplicados nas interações subsequentes.

Podem-se encontrar, no grupo, objetivos comuns, favoráveis e desfavoráveis. Contudo, o que prevalece é a eventual sedimentação de algo, podendo ser o próprio projeto utilizado interativamente para se chegar à sedimentação.

1.4. Auto-organização Secundária

Para que sejam possíveis processos de *auto-organização secundária*, é indispensável a existência de um ser ou sistema primariamente constituído por meio da *auto-organização primária*. A auto-organização secundária é um estágio posterior à auto-organização primária, visto que auto-organização secundária não é alcançada sem ter ocorrido inicialmente a primária.

Segundo Debrun (1996, p.13), a *auto-organização secundária* ocorre quando,

> num processo de aprendizagem (corporal, intelectual ou existencial), a interação se desenvolve entre as partes (mentais e/ou corporais) de um organismo - a distinção entre partes sendo então "semi-real" -, sob a direção hegemônica mas não dominante da "face-sujeito" desse organismo.

Debrun defende, portanto, que a auto-organização secundária pode ocorrer entre os elementos durante os processos de aprendizagem (ver capítulo 3), ao passo que a interação ocorre entre partes mentais e/ou corporais do organismo (grupo, sistema). Um elemento-guia traça o caminho do grupo, mas de uma maneira não dominante. Assim, em hipótese alguma deve existir um guia dominante que determine e ordene os passos do grupo.

1.5. Fundamentos de Sistemas Auto-organizados

A seguir, serão abordados os fundamentos teóricos de *sistemas auto-organizados*.

Ashby não aceita que sistemas não biológicos podem auto-organizar-se. Segundo esse autor (1962, p. 266), a afirmação de que um sistema é Auto-organizado dá lugar a dois diferentes significados. Primeiro o sistema

passaria – por si só - de um *estado desorganizado* para outro *organizado*; segundo, a organização adquirida pelo próprio sistema muda de *má organização* para *boa organização*.

Baseado na definição de Ashby (1962), o primeiro significado refere-se a sistemas que começam com as partes separadas em estado desorganizado (elementos independentes). As próprias partes (elementos), assim, agem de tal forma que elas próprias acabam mudando, formando conexões de algum tipo, compondo partes unidas (elementos dependentes). Ashby define isso como uma simples *auto-conexão* entre elementos.

Ashby diz que um sistema é *organizado* no sentido em que ele muda de *partes separadas* (elementos) para *partes unidas* (elementos dependentes), isto é, as partes separadas acabam por constituir um todo (ver Figura 2, seção 1.2).

Os passos seguintes apresentados por Ashby (1962, p. 267) mostram como um sistema passa de um estado desorganizado para organizado sendo o *Sistema* um Embrião do Sistema Nervoso

INÍCIO

- *Partes separadas (comportamento independente):* células que têm pouco ou nenhum conhecimento uma da outra;
- *Mudam e formam conexões de algum tipo (comportamento dependente):* crescimento dos dendritos e a formação de sinapses;
- *O sistema torna-se organizado, já que mudou de "partes separadas" para "partes unidas":* cada parte é intensamente afetada pelo comportamento de outras partes.

FIM.

Baseado no exemplo anterior, pode-se, dizer que esta é uma maneira de se formar um sistema organizado.

O segundo significado se refere ao sistema que muda de uma *má organização* para uma *boa organização*. Este é um caso de particular interesse, pois esses tipos de sistemas mudam a *si mesmos*, passando de mau a bom comportamento. É como se ocorresse um aperfeiçoamento entre as partes que se relacionaram durante o primeiro momento (auto-conexão). Para melhor compreensão do segundo significado, podem-se citar dois exemplos, segundo Ashby (1962, p. 267):

1° exemplo: considere-se uma criança que apresenta uma organização cerebral inicial que a faz aproximar-se do fogo. Em certo momento, uma mudança ocorre, uma nova organização cerebral surge, levando a criança a evitar o fogo.

Nesse exemplo, a criança que nunca viu ou nunca se aproximou do fogo e não o conhece, ao aproximar-se dele percebe que aquilo não lhe fará bem. Na próxima vez que a criança perceber a presença do fogo, ela irá evitá-lo.

2° exemplo: dois pilotos automáticos, de dois aviões, estão acoplados pelo erro de um feedback que não é corrigido; o erro se agrava e os aviões acabam por colidir. Aqui a organização é má. O sistema pode ser "auto-organizado" se uma mudança for feita automaticamente no feedback, mudando-o de positivo para negativo. Assim, o todo teria mudado de má para boa organização.

Neste exemplo, ocorreu uma má organização, visto que um feedback (retroalimentação) levou os aviões a se aproximarem ainda mais, ultrapassando o limite de aproximação permitido, o que resultou na colisão. O

correto seria que os aviões não excedessem o limite de aproximação determinado, evitando o acidente.

Analisando os dois exemplos, vê-se que no caso da criança, quanto mais ela se distancia do fogo, melhor será para ela. Neste caso, o sistema passou de má para boa organização. Já no exemplo do piloto automático, o ideal seria que a aproximação fosse realizada, mas mantendo-se certa distância, para que não ocorresse colisão. Como o sistema não conseguiu distinguir uma maior aproximação do que era permitido, pode-se dizer que ele possuía boa organização até certo ponto. Neste caso essa boa organização acabou por gerar uma má organização.

Portanto, deve-se enfatizar que não pode existir a suposição de que a organização desenvolvida será boa, analisando apenas uma situação particular. Uma *boa organização* pressupõe, em primeiro lugar, que se providenciem critérios para distinguir entre uma *organização boa* de uma *organização má* (Gonzales, 1996, pp. 284-285). Os critérios que nos possibilitarão fazer essa distinção são os *critérios de relevância*, a serem discutidos no final desse trabalho.

O passo seguinte será apresentar o *papel da informação* e do *ruído* junto aos sistemas auto-organizados.

1.6. O Papel da Informação e do Ruído em Sistemas Auto-organizados

Na seção 1.2 desta pesquisa, apontou-se como um de seus objetivos a investigação a respeito dos processos artificiais, compostos de

elementos distintos os quais acabam dando origem a sistemas auto-organizados. A interação entre esses elementos distintos comporão o sistema.

Segundo Gonzales (1996), acredita-se que é nesse ponto de interação que surge o *papel da informação*, porque as partes comunicantes, funcionando como fontes receptoras e emissoras, trocam informações de maneira simultânea.

Essa troca simultânea de informações entre as partes pode levar à formação de um sistema organizado, um todo. Um sistema pode ser constituído por um amontoado de partes que trocam informações, podendo evoluir de um estado com menor fluxo de informação entre suas partes, para um estado com maior fluxo de informação entre elas. Assim, de acordo com o fluxo (quantidade) de informação, o sistema passa de estados mais simples para estados mais complexos, ou de estados mais complexos para estados mais simples.

É importante salientar que nesse momento estamos tratando de sistemas fechados, uma vez que não se observam processos de auto-organização, ou seja, não há nenhuma interação com o ambiente externo. As regras de sistemas fechados são pré-definidas por instruções determinadas pelos programadores; portanto, não se trata de auto-organização e sim de hetero-organização.

Entretanto, supondo-se que esses sistemas não sejam fechados, e que sua condição de partida não possa ser uma ocorrência ou mudança de estado do sistema, o que, então, implica sua auto-organização? Onde surge seu início, sua condição de partida, seu evento inaugural?

Conforme Atlan (1992, p. 40),

> as únicas mudanças capazes de implicar a organização em si, e de não serem simplesmente mudanças de estado do sistema que façam parte de uma organização constante, têm que ser produzidas de fora do sistema.

Com essa afirmação, Atlan quer observar o fato de que ocorrerá auto-organização se as mudanças forem introduzidas via ambiente externo, provocadas pelo ruído, além, é claro, das mudanças que ocorrem internamente no sistema.

Segundo Shannon e Weaver (1949, p. 5), o ruído "é qualquer perturbação do canal que distorça ou, que de qualquer outro modo, massacre o sinal". Portanto, para Shannon e Weaver o ruído é o que danifica o sistema.

Já, Atlan (1992, pp. 38-40) não considera ruim a ocorrência de ruído no sistema. Além disso, sistemas auto-organizadores não se alimentam apenas da ordem, mas também podem utilizar o ruído como outra fonte de alimentação. Complementa dizendo que a ocorrência de auto-organização via ruído, pode ser possível de duas maneiras diferentes: via um *programa preciso*, injetado no sistema por um programador, que determinaria as mudanças sucessivas da organização funcional, ou por meio de mudanças sucessivas *determinadas pelo exterior*, fatores aleatórios para os quais é impossível estabelecer qualquer lei que prefigure uma organização. Atlan conclui dizendo que esse é o momento em que se pode falar de auto-organização.

Para Shannon e Weaver, o ruído é visto como um destruidor de sinais, enquanto que para Atlan ele pode ser uma fonte de alimentação, transformada num fator de organização.

Quanto à condição de partida ou evento inaugural, Silva (1996, pp. 117-127) relata que é nesse momento que surge o *papel do ruído*. O ruído é que dá início ao sistema, criando o evento inaugural externo.

Segundo Silva (1996, pp. 117-127), existem duas maneiras por meio das quais o ruído pode influenciar os sistemas auto-organizados. A primeira é a de *catalisador*, cujo objetivo é colocar em ação os processos internos do sistema, levando-o a se estabelecer e aumentar a eficiência dos canais de comunicação quanto à interação entre as partes do sistema. A segunda é a de *afrouxador dos canais de comunicação* entre as partes do sistema, deixando entrar mais informações, as quais podem, dependendo das condições, vir a gerar complexidade organizacional. A esse processo Silva dá o nome de *ruído informacional*.

1.7. Graus da Auto-organização

Nesta seção, serão apresentados os graus nos quais os processos de auto-organização podem ou não ocorrer.

De acordo com Debrun (1996, pp.30-31), pode-se dividir a ocorrência ou não de auto-organização em cinco graus: os processos que ficam excluídos da auto-organização; a auto-organização de "grau zero"; a de "grau 1"; a de "grau 2"; e os processos de auto-organização para sedimentação do próprio ser em questão.

A seguir, serão discutidos cada um dos cinco graus acima mencionados.

- **Processos que ficam excluídos da auto-organização**

Os processos que envolvem automação mecânica, como, por exemplo, o funcionamento de relógios, autoramas, etc, são processos que ficam excluídos da auto-organização, visto que esses processos percorrem caminhos rígidos, limitados, sem alternativas ou opções de escolha, isto é, não têm mecanismos que possibilitem ajustes e aprimoramentos.

- **"Grau zero" da auto-organização**

São processos cujos caminhos não são rígidos como os do grau mencionado anteriormente. Aqui há alternativas de desvios, mais ou menos imprevisíveis, que sofrem uma correção quase que mecânica, permitindo a reprodução quase correta da situação já referida. A idéia aqui pende para a existência de um *quase ser*, no qual podem ocorrer ajustes e aprimoramentos quase que mecânicos. Portanto, pode-se chamar isso de uma "pré auto-organização".

- **"Grau 1" da auto-organização**

Pode-se considerar como processos de *grau 1* da auto-organização, o autofuncionamento cibernético, visto que esses processos podem dar respostas flexíveis por meio da evolução de uma situação. Além disso, esses processos são orientados por normas a serem atingidas ou até mesmo mantidas, implicando o uso de outros meios ou caminhos para alcançar determinado objetivo, que, por sua vez, ao menos nas linhas mais gerais, não podem ser redefinidas pelo próprio sistema.

- **"Grau 2" da auto-organização**

Como *grau 2* da auto-organização, podem-se citar sistemas complexos, em particular sistemas humanos, já cristalizados, nos quais se podem redefinir seus eventuais objetivos, mas não a si próprio. Ou seja, a redefinição de seus objetivos pode ocorrer inúmeras vezes até que se alcance um resultado de melhor qualidade quanto ao objetivo inicialmente proposto; o que, em hipótese alguma, pode ser redefinido, já que é o próprio ser em questão.

- **Sedimentação de si próprio**

Neste caso, haverá processos de auto-organização quando um sistema consegue "ser a própria gênese do seu ser", ou se redefinir parcialmente por aprendizagem e/ou incorporação de elementos externos, alcançando assim, novos patamares de atividade. Todavia, não pode existir a capacidade de se produzir de ponta a ponta a partir de uma lei de construção ou de um algoritmo.

Em resumo, foram apresentados, neste capítulo, os fundamentos da *Teoria da Auto-organização*, passando-se obrigatoriamente pelos conceitos de *Sistemas*, de "Organização, de "Auto" e de *processos de auto-organização*, bem como, pelas noções de *Condição de Partida*, *Elementos Distintos*, *Encontro e Interação*, *Sujeito Onipotente* e *Forma Organizada*, já que são estas as noções centrais subjacentes aos processos de auto-organização. Foram vistos também os dois tipos de auto-organização, a *primária* e a *secundária*, bem como os fundamentos teóricos associados aos *sistemas auto-organizados*. Realizou-se ainda, um estudo sobre o *papel da informação* e do *ruído* nos sistemas auto-organizados. Por último, apresentou-se os *graus da auto-organização* que podem ou não ocorrer em processos de auto-organização

unesp

O próximo capítulo, tratará de apresentar os fundamentos dos modelos de *Redes Neurais Artificiais*, indicando os elementos fundamentais para a sua composição.

2

Redes Neurais Artificiais

Este capítulo foi dividido em cinco seções e, como o anterior, é considerado primordial para que se atinja o objetivo final deste trabalho: *constatar se processos de auto-organização ocorrem nos modelos de redes neurais artificiais*.

Na primeira seção, apresentamos brevemente os fundamentos das *Redes Neurais Artificiais* e em que área de trabalho atual elas podem ser utilizadas para apoio à tomada de decisão.

A segunda seção refere-se à motivação das *pesquisas a respeito das redes neurais artificiais*, ou seja, a composição do *cérebro humano*.

Aborda-se na terceira seção, com detalhes, o elemento central das redes neurais artificiais: o *neurônio artificial* ou *unidade de processamento*.

Na seção subsequente, destacam-se os conceitos básicos sobre as *simulações de formação de padrões informacionais*, bem como as duas vias pelas quais uma rede neural artificial pode exercer essa simulação: *via supervisor* e *via meio ambiente externo*.

A quinta seção refere-se às *principais arquiteturas* que as redes neurais artificiais podem assumir, visto que é por meio delas que se pode restringir o tipo de problema envolvido por uma rede neural artificial.

2.1. Fundamentos relativos às Redes Neurais Artificiais

Conforme Marar (1997, p. 65), as redes neurais artificiais são "sistemas paralelos distribuídos, formados por *unidades de processamento*, que tentam imitar ou simular as mais elementares funções do *neurônio biológico*".

Apesar de a ciência ainda não conhecer plenamente o funcionamento neurofisiológico cerebral, unidades de processamento ou unidade comportamental básica, na maioria das vezes, são baseadas nas características do funcionamento cerebral.

Segundo Haykin (2001, p. 28), uma rede neural artificial é

> um processador maciçamente paralelo distribuído, constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso.

Ainda, segundo Haykin (2001, p. 28), as redes neurais artificiais são também referidas na literatura como *neurocomputadores* ou *redes conexionistas*. Uma rede neural artificial se assemelha ao cérebro em dois aspectos: no primeiro, a rede neural adquire conhecimento a partir de seu ambiente, mediante um processo de aprendizagem; no segundo, para armazenar o conhecimento adquirido, a rede neural utiliza-se das forças de conexões entre as unidades de processamento, conhecidas como pesos sinápticos.

Vale acrescentar que, além das unidades de processamento, as redes neurais artificiais possuem uma arquitetura e um método de aprendizagem (algoritmo). As arquiteturas, apresentadas na seção 2.5, também conhecidas por topologias, definem as estruturas utilizadas para as conexões entre as unidades de

processamento. Essas estruturas podem ser definidas pelo número de camadas assumidas pela rede neural artificial, dependendo do problema a ser tratado. Os métodos de aprendizagem, via algoritmo, que realizam os ajustes das conexões entre as unidades de processamento, serão tratados no capítulo 3, seção 3.2.

Atualmente, podem-se utilizar as redes neurais artificiais como ferramentas de apoio às tomadas de decisões, em diversas áreas profissionais e científicas: áreas militares, geográficas, medicinais, econômicas, etc. Como exemplos de aplicações, podemos citar: previsões meteorológicas e oceanográficas, processamento de sinais complexos, tal como radar, diagnósticos médicos, reconhecimento de voz e imagens, classificação de sinais, etc.

Considerando-se as formas de utilização apresentadas para as redes neurais artificiais nas áreas profissionais e científicas, pode-se dizer que as redes neurais artificiais poderão ter, futuramente, uma grande função em novas descobertas tecnológicas.

2.2. O que motivou as pesquisas sobre as Redes Neurais Artificiais

A principal fonte de inspiração para as redes neurais artificiais foram os neurônios existentes no cérebro humano, constituindo-se em seus elementos fundamentais e em unidades básicas de processamento. O neurônio artificial é o elemento básico das redes neurais artificiais. Como o objetivo das redes neurais artificiais é simular o neurônio biológico, por meio de uma unidade de processamento (neurônio artificial), o cérebro humano se tornou o grande motivador das redes neurais artificiais.

Segundo Haykin (2001), estima-se que haja aproximadamente 10 bilhões de neurônios no córtex humano e 60 trilhões de sinapses ou conexões.

Um bom exemplo dessa motivação é o modelo proposto por McCulloch e Pitts em 1943, que propõe elementos computacionais abstraídos das propriedades fisiológicas dos neurônios biológicos, formadores do cérebro humano e de suas conexões.

O neurônio biológico é delimitado por uma membrana celular que, além da função biológica normal, possui propriedades essenciais para a sua capacidade de processar e transmitir informações. Cada um, dentre a grande quantidade de neurônios do cérebro humano, processa informações e se comunica com outros milhares de maneira contínua e paralela. Cada um consegue receber e transmitir informações ao mesmo tempo e em paralelo com outros milhares de neurônios.

A base das investigações sobre redes neurais artificiais esteve, inicialmente, vinculada ao neurônio biológico, isto é, à sua estrutura individual, às suas formas de conexões (formas de processar informações e de se comunicar) e ao comportamento de uma forma conjunta de milhares de neurônios biológicos, ou seja, de uma rede neural biológica.

Segundo Braga (1998, p. 4), o cérebro humano é

> responsável por nossas emoções, pensamentos, percepções e cognições, execução de funções senso-motoras e autônomas. Além disso, sua rede de neurônios tem a capacidade de reconhecer padrões e relacioná-los, usar e armazenar conhecimento por experiência, além de interpretar observações. Mesmo com tudo isso, o funcionamento das redes biológicas ainda não foi totalmente desvendado pelo homem, mas é nesta estrutura fisiológica básica de redes de neurônios naturais que se baseiam as redes neurais artificiais.

O objetivo das redes neurais artificiais é simular tarefas intelectuais complexas, realizadas pelo cérebro humano, por meio de sistemas computacionais. Dentre as tarefas intelectuais complexas estão: a resolução de problemas, reconhecimento e classificação de padrões informacionais, processos indutivos e dedutivos, ou seja, tudo que o cérebro humano possa realizar, como por exemplo somar 1 + 1.

As redes neurais artificiais se inspiraram no cérebro humano, no qual o elemento biológico principal é o neurônio, em sua maneira de se organizar, em sua capacidade de executar tarefas computacionais complexas.

Portanto, pode-se considerar que é por meio das redes neurais artificiais que se tentam simular certas funções do cérebro humano, baseando-se em unidades de processamento chamadas, neurônios artificiais, que recebem, processam e transmitem informações e se comunicam paralela e simultaneamente. Essas unidades de processamento se relacionam por meio de conexões, recebendo e transmitindo informações.

2.3. Unidade de Processamento (Neurônio Artificial)

Esta seção enfocará as unidades básicas formadoras de redes neurais artificiais, denominadas *unidades de processamento*. Essas unidades também são conhecidas como UNS's, *Unidades Neurônio-Símile*.

Haykin (2001, p. 36) considera que a unidade de processamento é o elemento fundamental para a operação de uma rede neural artificial.

Segundo Marar (1997, p. 69), uma unidade de processamento "é a unidade que computa a função limiar, que coleta as entradas e produz uma saída, apenas se a soma das entradas exceder um valor limiar interno".

Detalhando a definição dada por Marar: uma unidade de processamento é composta por *terminais de entrada*, ou seja, os meios pelos quais os estímulos de entrada são inseridos na rede neural; por uma *função soma*, que computa os estímulos de entradas multiplicados pelos seus respectivos pesos; e por um *terminal de saída*, por meio do qual a rede neural dá o resultado proveniente da função soma.

Apresentada de uma maneira mais técnica, segundo Braga (1998, pp.7-8), uma unidade de processamento é constituída,

> por meio a uma função matemática composta de terminais de entrada (x_1, x_2, ... , x_n), que representam os dendritos, e por apenas um terminal de saída (y), representando o axônio. Também fazem parte desta função os pesos acoplados, como por exemplo (w_1, w_2, ... , w_n), cujos valores podem ser positivos ou negativos, dependendo de as sinapses correspondentes serem inibitórias ou excitatórias.

Na definição anterior, Braga utilizou os termos *dendritos*, *axônios*, *resistores* e *sinapses* pertinentes ao neurônio biológico, procurando fazer analogia entre a definição da unidade de processamento com a constituição do neurônio biológico. Para ilustrar esta analogia pode-se citar a unidade de processamento proposta por McCulloch e Pitts em 1943, conhecida por unidade de processamento MCP (ver Figura 7).

A seguir detalhar-se-á cada um dos elementos que compõem uma unidade de processamento

- **Dendritos ou Terminais de Entrada, ou simplesmente Entrada**

Segundo Haykin (2001) os dendritos são filamentos receptivos que possuem uma superfície irregular com muitas ramificações (sendo assim chamados pela sua semelhança com uma árvore). É por meio dos *dendritos* que os estímulos ou impulsos de entrada são recebidos pelas unidades de processamento. Não existe um limite fixo de estímulos de entrada, isto é, as unidades de processamento podem receber entre 1 a n estímulos de entrada. Eles podem ser comparados a estímulos do neurônio biológico. Todos os estímulos são enviados à unidade de processamento simultaneamente, ou seja, se uma unidade de processamento possuir três estímulos de entrada, os sinais dos três estímulos deverão chegar até o núcleo da unidade de processamento paralela e simultaneamente.

- **Axônio ou Terminal de Saída ou simplesmente Saída**

Segundo Haykin (2001), os axônios são filamentos de transmissão que possuem uma superfície lisa, com poucas ramificações e de grande comprimento. Os estímulos de saída são propagados, via axônio pré-sináptico, para a unidade de processamento pós-sináptica. Individualmente, uma unidade de processamento possui um único axônio, embora este possa apresentar algumas ramificações.

- **Sinapses ou Conexões Sinápticas (elos de conexão)**

Segundo Haykin (2000), as sinapses são unidades estruturais e funcionais que mediam as instruções entre as unidades de processamento pré-sináptico (axônio) e as unidades de processamento pós-sinápticos (dendritos).

Essas duas partes, pré e pós-sinápticas, são separadas pela *fenda sináptica*. Conforme Kovács (1997, p. 21), a propagação da informação, vinda da unidade pré-sináptica em direção à pós-sináptica, é realizada via *neurotransmissores*, por meio da fenda sináptica. A grosso modo, pode-se dizer que as sinapses são regiões nas quais duas unidades entram em contato (fenda sináptica), transmitindo e recebendo estímulos.

- **Neurotransmissores ou Pesos**

Os *neurotransmissores* são valores acoplados aos estímulos de entrada. Os valores podem ser positivos (sinapses excitatórias) ou negativos (sinapses inibitórias). Esses valores também são conhecidos como pesos sinápticos, por serem valores transmitidos via sinapses. Os valores são representantes do grau de importância, que determinado estímulo de entrada possui, em relação a uma determinada unidade. Esses valores são multiplicados pelos estímulos de entrada apresentados à unidade de processamento e somados. Esse resultado poderá excitar ou inibir a unidade de processamento.

- **Função Soma**

A *função soma* tem por atribuição fazer a somatória da multiplicação dos estímulos de entrada (*inputs*) recebidos de outras unidades por seus respectivos pesos, resultando assim em um valor que será comparado ao limite máximo fixado pela rede neural, *threshold*. Se esse valor não ultrapassar o *threshold*, ele será transferido para a próxima unidade de processamento, via função de ativação. Caso contrário, os pesos são reajustados a fim de atualizar esse valor, até, se possível, torná-lo um valor transferível.

- **Função de Ativação**

A *função de ativação* destina-se ativar a unidade de processamento posterior por meio do valor proveniente da função soma. Esta função é também conhecida como limiar lógico ou *threshold*. É ela que define o *threshold* utilizado pela rede, na comparação da função soma, para se saber se a próxima unidade de processamento será ou não ativada.

Conforme Lippmann (1987), as funções de ativação Limite Ríspido (*Hard Limiter*), Limite Lógico (*Threshold Logic*) e Sigmóide (*Sigmoid*) são as mais conhecidas.

Na função Limite Ríspido ou Função de Limiar, segundo Haykin (2001), a saída (y) de uma unidade de processamento assume o valor 1, se o valor retornado pela função soma não for negativo; e assumirá o valor -1 caso o valor retornado pela função soma seja negativo.

A Figura 4, ilustra esse tipo de função.

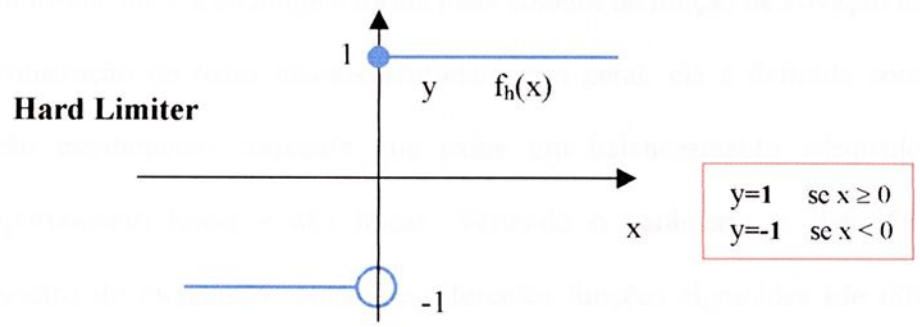

Figura 4. Função Limite Ríspido (*Hard Limiter*), inspirada em Haykin (2001, p. 39) e Lippmann (1987, p. 14).

Também conhecida por Função Rampa (Haykin, 2001), na Função Limite Lógico, a saída (y) de uma unidade de processamento assume o

valor 1, se o valor retornado pela função soma for maior ou igual a 0,5; assumirá o valor de x, caso o valor retornado seja maior que -0.5 e menor que 0,5; e assumirá o valor -1, se o valor retornado for menor ou igual a -0,5.

A figura a seguir, ilustra esse tipo de função.

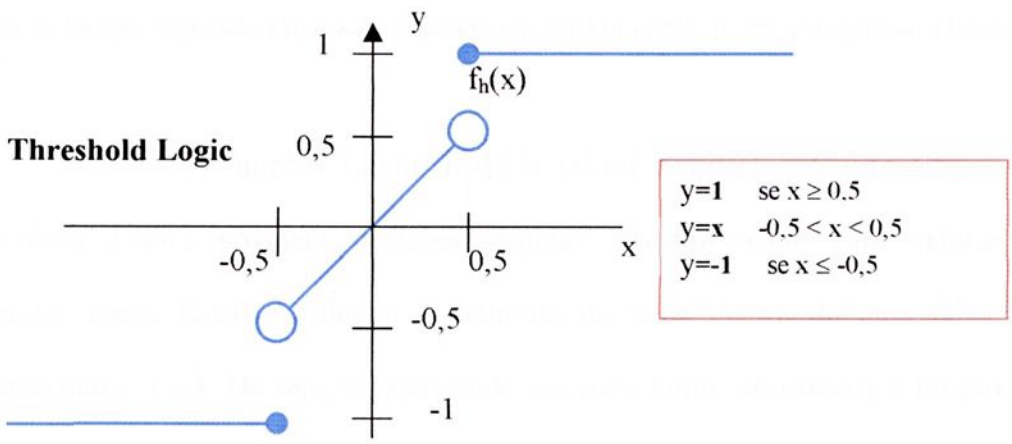

Figura 5. Função Limite Lógico (*Threshold Logic*), inspirada em Haykin (2001, p. 39) e Lippmann (1987, p. 14).

Segundo Haykin (2001, p. 39), a função sigmóide, cujo gráfico tem a forma de *S*, é de longe a forma mais comum de função de ativação utilizada na construção de redes neurais artificiais. Em geral, ela é definida como uma função estritamente crescente que exibe um balanceamento adequado entre comportamento linear e não linear. Variando o parâmetro *c*, definido como parâmetro de inclinação, obtém-se diferentes funções sigmóides (de diferentes inclinações). A saída dessa função assume valores contínuos entre o intervalo de 0 e 1.

A Figura 6 ilustra esse tipo de função.

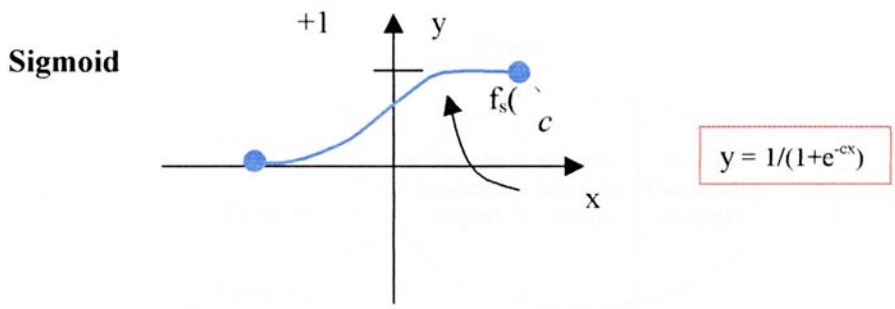

Figura 6. Função Sigmoide (*Sigmóid*), inspirada em Haykin (2001, p. 39) e Lippmann (1987, p.14).

As funções Limite Ríspido (*Hard Limiter*) e Limite Lógico (*Threshold Logic*) possuem decisões ríspidas, principalmente para valores extremos. Essas funções refletem o estímulo de saída dentro de uma faixa, digamos entre -1 e 1. Ou seja, ultrapassando um certo limite (*threshold*), a função dispara o valor -1 ou 1 ou x, dependendo da função e da forma que será utilizada.

A Função Sigmóide (*Sigmoid*) é utilizada para decisões em que os estímulos de saída limites são disparados quando existe uma saturação muito alta do valor de ativação. Ela produzirá um valor 0 de estímulo de saída somente quando o valor passado da função soma tiver um grande valor negativo; e produzirá o valor de estímulo de saída 1, quando a soma tiver um grande valor positivo, ou seja, ela é gradual, tende a zero de um lado e a 1 do outro lado.

Após as definições sobre unidade de processamento, ou funções de ativação, pode-se, ilustrar uma composição individual de uma unidade de processamento.

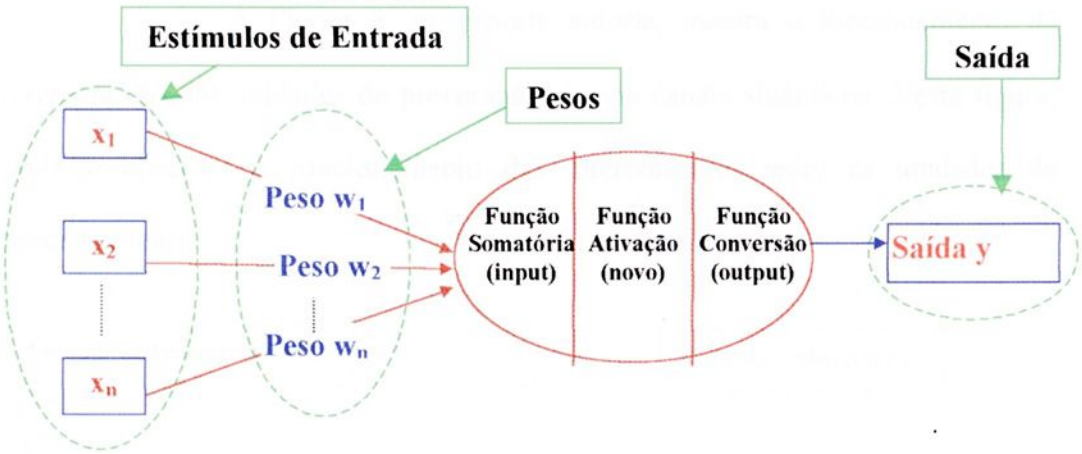

Figura 7. Composição de uma unidade de processamento (neurônio), inspirada em Braga (1998, p.8).

Conforme Lippman (1987, p.14), a unidade de processamento, ilustrada na Figura 7, entra em funcionamento por meio da função soma descrita a seguir:

$$y = \sum_{i=1}^{n} WiXi - \theta$$

y = estímulo de saída

n = é o número de estímulos de entrada da unidade de processamento;

Xi = são os estímulos de entrada;

Wi = são os pesos acoplados ou associados aos estímulos de entrada Xi;

θ = limiar fixo (*threshold*)

Analisando a Figura 7, nota-se que além dos estímulos de entrada definidos como X_1, X_2,...,X_n, pesos definidos como W_1, W_2,...,W_n e do estímulo de saída definido como y, tem-se também as funções soma e ativação que completam a composição da unidade de processamento.

A Figura 8, de própria autoria, mostra o funcionamento da interconexão das unidades de processamento via canais sinápticos. Nesta figura, pode-se observar o funcionamento das interconexões entre as unidades de processamento.

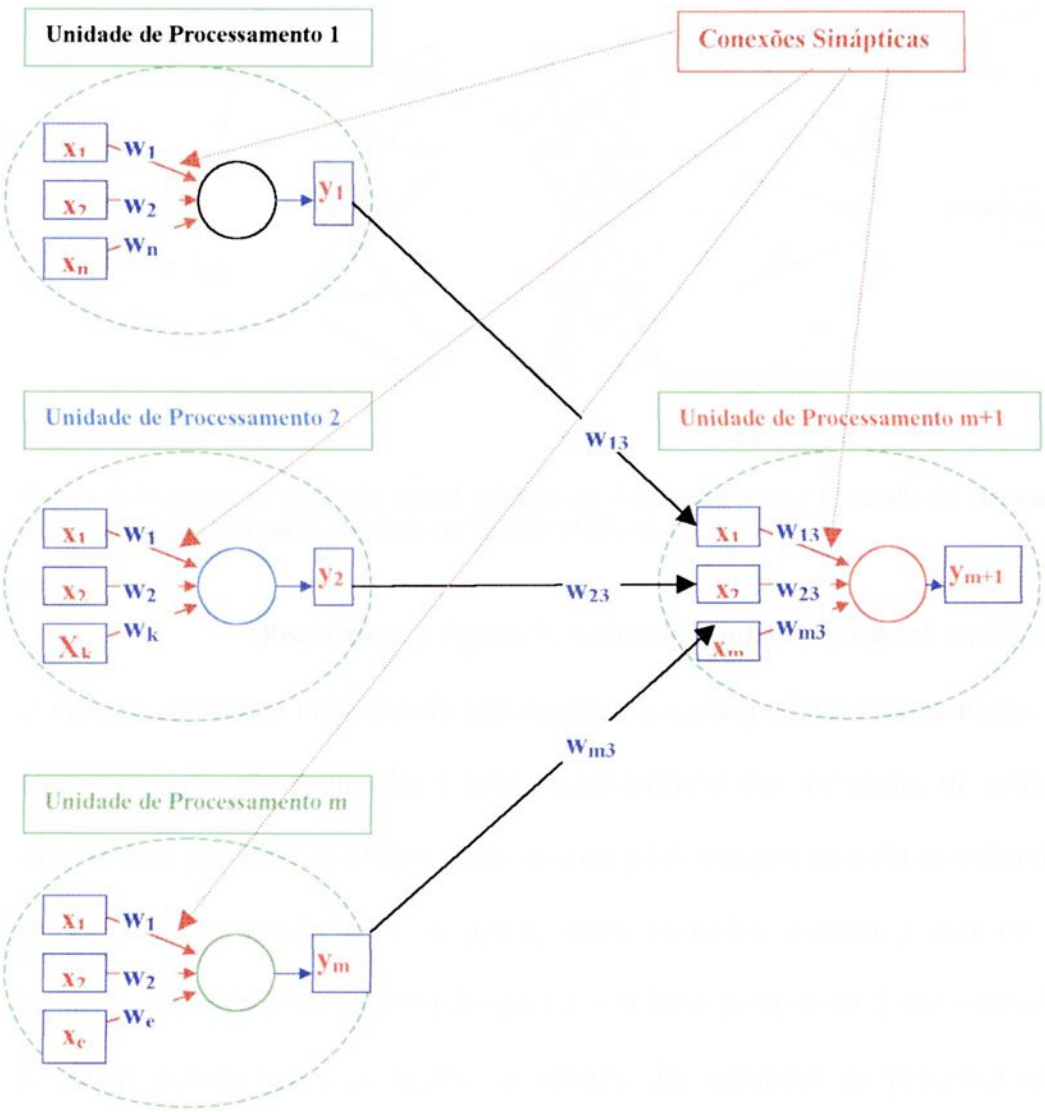

Figura 8. Composição detalhada das interconexões entre várias unidades de processamento.

Neste momento, considera-se que seja adequada a ilustração de uma rede neural artificial, uma vez que já se tem a ilustração individual de uma

unidade de processamento (Figura 7), bem como de várias unidades interconectadas (Figura 8). A Figura 9 ilustra, assim, uma rede neural artificial por completo.

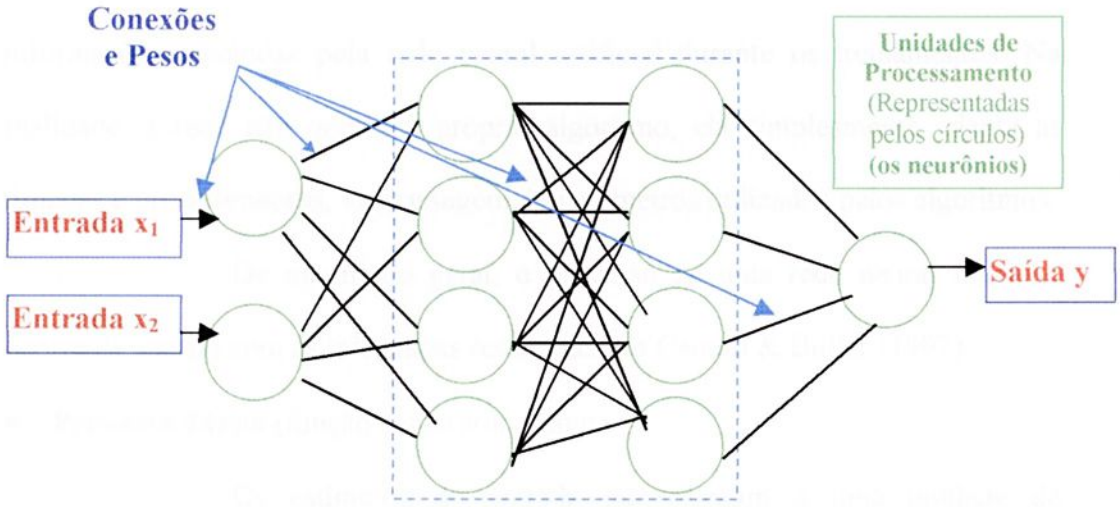

Figura 9. Exemplo de uma rede neural artificial de 4 camadas, sendo 1 camada de entrada, 2 intermediárias e 1 de saída, inspirada em Kovács (1996, p. 40).

Observando a Figura 9, vemos que uma rede neural artificial é composta por várias unidades de processamento, cada qual descrita na Figura 7. Neste exemplo, são fornecidos à rede neural artificial dois estímulos de entrada para as duas primeiras unidades. Após as duas primeiras processarem os estímulos de entrada, de acordo com os pesos, esses estímulos passam a compor os estímulos de saída e são transmitidas para as unidades posteriores. Esses estímulos de saída tornam-se os estímulos de entrada das unidades de processamento posteriores. E assim sucessivamente, até que a rede encontre a última camada de unidades de processamento, resultando o estímulo de saída final da rede.

Na Figura 9, todas as unidades anteriores estão totalmente conectadas a todas as unidades posteriores.

Para que a rede neural artificial entre em funcionamento são necessárias informações que incorporem regras para a aprendizagem. Essas regras são incorporadas ao algoritmo, desenvolvido por um programador, via procedimentos e funções, no qual poderão sofrer adaptações e ajustes via informações recebidas pela rede neural artificial durante os treinamentos. Na realidade, a rede não gera seu próprio algoritmo, ela simplesmente adapta as funções e procedimentos, via passagem de parâmetros utilizados pelos algoritmos.

De um modo geral, o processo de uma rede neural artificial ocorre de acordo com as três etapas resumidas por Caudill & Butler (1992):

- **Primeira Etapa** (função somatória - input)

Os estímulos de entrada que chegam a uma unidade de processamento são somados obtendo-se o valor total, ou "*net input*", normalmente calculado por meio da função limiar, com uma pequena mudança no que diz respeito à unidade de processamento que se conectará à próxima unidade. Portanto, acrescenta-se o *j*, que indica qual é a unidade de processamento anterior que se conectará a próxima unidade. A função da página 45 é utilizada para o funcionamento de uma única unidade. Já função seguinte destina-se ao funcionamento de uma rede, ou seja, para mais de uma unidade de processamento em conexão.

$$y_j = \sum_{i=1}^{n} Wij\,Xij - \theta$$

$$y_j = W_{1j}.X_{1j} + W_{2j}.X_{2j} + W_{3j}.X_{3j} + \ldots + W_{nj}.X_{nj} - \theta$$

cm 1 2 3 4 5 6 7 unesp 1 12 13 14 15 16 17 18

i e j = são índices das unidades de processamento;

yj = valor da soma de todos os estímulos de entrada recebidos pela unidade i;

Xij = é o valor do estímulo proveniente da unidade i que chega até a unidade j;

Wij = é o peso proveniente da unidade i que chega até a unidade j;

n = é o último elemento da seqüência, a quantidade de entradas.

θ = limiar fixo (*threshold*)

Esta função leva y_j a obter a informação sobre o comportamento geral das unidades de processamento.

- **Segunda Etapa** (função ativação)

Consiste em transformar os estímulos de entrada em um grau de ativação da unidade de processamento. Essa transformação é comum para todas as unidades que terão valores limite de atuação. Geralmente tal etapa é descrita nas redes neurais por meio de uma função sigmóide:

$$f(i) = \frac{1}{1 + e^{-ci}}$$

Nessa função, o valor mínimo tende para 0 (zero) e o valor máximo tende para 1 (um).

- **Terceira Etapa** (função conversão - output)

Consiste em converter o grau da ativação das unidades de processamento para um estímulo de saída (y) destas unidades. Ele é geralmente estabelecido de acordo com a função utilizada pela rede. Neste caso utilizamos a função sigmóide:

$$y_j = \begin{cases} 1, f(i) > \theta, \\ 0, f(i) \leq \theta, \end{cases}$$

em que:

θ = é um valor limite fixado (*threshold*)

y_i = é o valor do estímulo de saída da unidade *i*

Assim, esta expressão define que o estímulo de saída de uma unidade de processamento sera igual ao seu valor de ativação sempre que ele exceder o valor limite θ, e será nulo quando este valor for menor que o limite θ.

A Figura 10, de própria autoria, ilustra uma unidade de processamento mediante estímulos de entrada, passando pelas três funções descritas anteriormente e finalizando com uma saída proveniente dos estímulos de entrada.

Figura 10. Exemplo de uma unidade de processamento ou uma *unidade neurônio-símile* (UNS).

Nessas três etapas, os estímulos de entrada internos e aqueles provenientes do meio ambiente externo são inseridos nas unidades de processamento, as quais constituirão os modelos de redes neurais artificiais para a simulação de determinados comportamentos relacionados à atividade de reconhecimento de padrões informacionais (representáveis).

Segundo Kovács (1996), uma rede neural artificial "não é um computador como concebemos hoje no senso comum, uma vez que ela não é programada como um computador".

Kovács (1996) complementa dizendo que

> os programas computacionais, estruturados na forma de sistemas fechados não possuem autonomia, uma vez que eles são pré-elaborados pelo programador, sem qualquer possibilidade de alteração significativa das suas regras de funcionamento. Ao contrário, as redes neurais artificiais contêm uma grande capacidade de se auto-organizar, dados que incorporam regras para a aprendizagem, gerando assim seu próprio algoritmo. Isso possibilita que as conexões ponderadas se auto-ajustem entre os elementos de processamento, mediante sessões de treinamento ou em respostas a uma nova entrada.

Kovács considera, portanto, que, quando se trata de redes neurais artificiais, trabalha-se com regras que são incorporadas nestas redes. Essas regras, segundo o auto, são incorporadas pela rede de acordo com as informações por ela recebidas. Isso leva a rede a adaptar o algoritmo, ou seja, a ajustar as regras presentes nos algoritmos, possuindo assim, uma certa autonomia sobre qualquer regra a ser elaborada.

Portanto, conforme Kovács, pode-se dizer que as redes neurais artificiais aprendem a resolver um problema. Já em programas computacionais as regras são pré-elaboradas pelo programador, isto é, não há qualquer interferência nas regras de funcionamento. Portanto, aqueles programas são considerados sistemas fechados, pois não possuem autonomia.

Quanto ao fato da rede neural artificial gerar seu próprio algoritmo, na realidade, a rede neural artificial apenas adapta o algoritmo via parâmetros, e não o cria (gera) como cita Kovács (1996).

2.4. Conceitos básicos sobre Formação de Padrões Informacionais

Nesta seção, será discutida a maneira como se dá a simulação das redes neurais artificiais e por quais propriedades as redes neurais artificiais simulam a atividade que possibilita a formação de padrões informacionais. Ou seja, a produção de algo proveniente do meio ambiente externo ou mesmo fornecido por um supervisor.

Segundo Gonzales (1996, p. 276), são as formações de padrões informacionais

> as principais propriedades dos modelos de redes neurais artificiais que simulam a atividade de formação de padrões informacionais, isto é, dados os sinais desejados (inputs) ou somente fluxos informacionais extraídos do meio ambiente, espera-se que a rede neural artificial consiga reproduzir uma saída desejada dada ou produzir uma saída proveniente do fluxo informacional extraído do meio ambiente.

As figuras a seguir, de própria autoria, ilustram de acordo com a definição de Gonzales, os dois meios pelos quais uma rede neural artificial pode exercer a simulação de formação de padrões informacionais.

A Figura 11(a) exerce a simulação com sinais desejados, definidos pelo supervisor da rede neural artificial. A Figura 11(b) apresenta a simulação com sinais oriundos do meio ambiente externo, sem a presença de um supervisor.

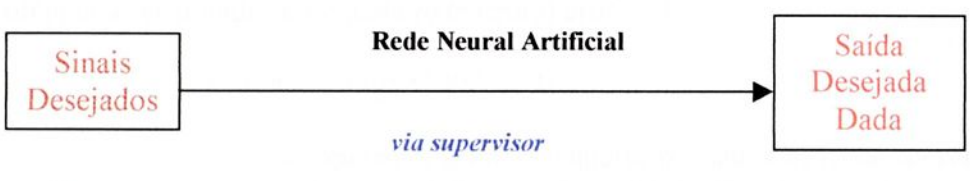

Figura 11(a). Modelo de Simulação com sinais fornecidos pelo supervisor.

Figura 11(b). Modelo de Simulação com sinais provindos do meio ambiente.

Conforme Gonzales (1996, p. 276), nas redes neurais artificiais

> a formação e o aperfeiçoamento de padrões podem ser caracterizados por meio de conexões que se estabelecem entre conjuntos de unidades de processamento, conhecidas também como unidade neurônio-símile (UNS's), as quais compõem sua arquitetura.

As unidades de processamento, como se viu anteriormente na seção 2.3, são aquelas que realizam o processo de receber e transmitir estímulos por meio das suas conexões, as *sinapses*. Tal processo pode variar de acordo com os diferentes modelos de redes neurais artificiais existentes. A partir desses modelos serão discutidos dois tipos básicos de aprendizagem: (1) *aprendizagem supervisionada* e (2) *aprendizagem não supervisionada* ou *por auto-organização*.

2.5. Principais Arquiteturas das Redes Neurais Artificiais

A concepção da arquitetura de uma rede neural artificial é muito importante, já que é mediante as arquiteturas que se pode restringir o tipo de problema a ser tratado por aquela rede neural artificial.

Segundo Braga (1998, p. 10),

> a arquitetura geral é composta por um número de camadas da rede; por um número de unidades de processamento em cada camada; por um tipo de conexão entre as unidades de processamento e pelo tipo de conexão utilizado.

Braga define as redes neurais artificiais de *camada única* como aquelas nas quais só existe uma unidade de processamento entre quaisquer estímulos de entrada e quaisquer estímulos de saída. As Figuras 12.a e 12.d ilustram esse tipo de arquitetura. O autor da o nome de redes neurais artificiais de *múltiplas camadas* aquelas nas quais possuem mais de uma camada de processamento entre qualquer estímulo de entrada e a saída da rede neural artificial. Esses tipos de arquiteturas estão ilustrados nas Figuras 12.b e 12.c.

Classificando as arquiteturas por tipos de conexões, Braga define como redes neurais artificiais *acíclicas* (*feedforward*) (alimentadas adiante) aquelas em que o estímulo de saída de uma unidade de processamento na *i*-ésima camada da rede neural artificial não pode ser usada como estímulo de entrada de unidades de processamento em camadas de índice menor ou igual a *i* (ver Figuras 12.a e 12.b). Como *cíclicas* (*feedback*) (alimentadas adiante com pelo menos um laço de realimentação) designa aquelas em que o estímulo da saída de uma unidade de processamento na *i*-ésima camada da rede neural artificial pode ser usado como estímulo de entrada de uma unidade de processamento em camadas de índice menor ou igual a *i* (ver Figuras 12.c e 12.d).

Por último, Braga classifica as arquiteturas quanto à sua conectividade. As redes neurais artificiais são denominadas *parcialmente conectadas*, quando nem todos os estímulos de entrada estão conectados com todas unidades de processamento. Pode-se, então, considerar as Figuras 12.b e 12.c como redes neurais artificiais, parcialmente conectadas. Analisando a Figura 12.b, verifica-se que os estímulos de entrada x_1, x_2, x_3 só estão conectados à

unidade de processamento 1; e os estímulos de entrada x_4, x_5, x_6 só estão conectados à unidade de processamento 2. No entanto, os valores dos estímulos de entrada acabam encontrando-se na unidade de processamento 3 e compõem o estímulo de saída y_1. Braga, concebe as redes neurais artificiais *totalmente conectadas* quando todos os estímulos de entrada estão atrelados a todas as unidades de processamento. As Figuras 12.a e 12.d são redes neurais artificiais totalmente conectadas.

Analisando a Figura 12.a, verifica-se que os estímulos de entrada x_1, x_2, x_3 estão conectados a todas as três unidades de processamento. Com isso, cada unidade de processamento pode trabalhar com os valores de todos os estímulos de entrada simultaneamente, gerando, cada unidade de processamento, seu estímulo de saída, que servirá de estímulo de entrada para as unidades de processamento posteriores.

A seguir são apresentados alguns exemplos de arquiteturas de redes neurais artificiais baseadas em Braga (1998, p. 11):

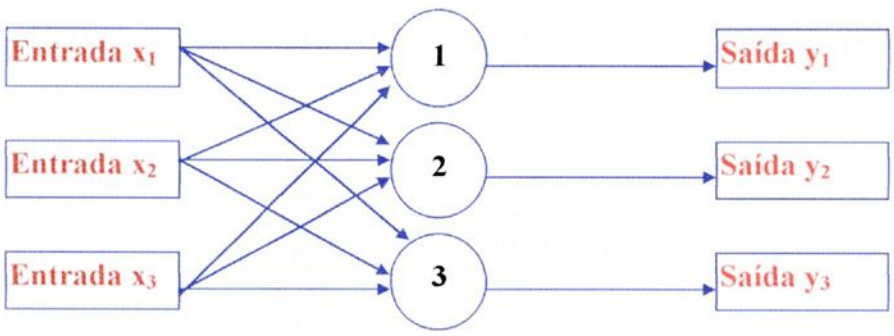

Figura 12.a. Camada Única -Acíclica - Totalmente Conectada, inspirada em Braga (1998, p. 11).

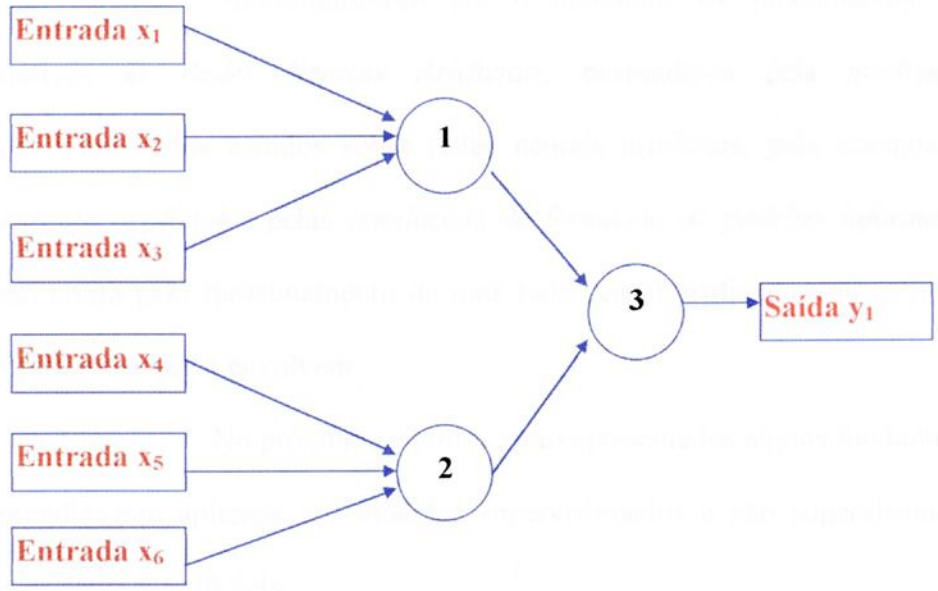

Figura 12.b. Múltiplas Camadas – Acíclica - Parcialmente Conectada, inspirada em Braga (1998, p. 11).

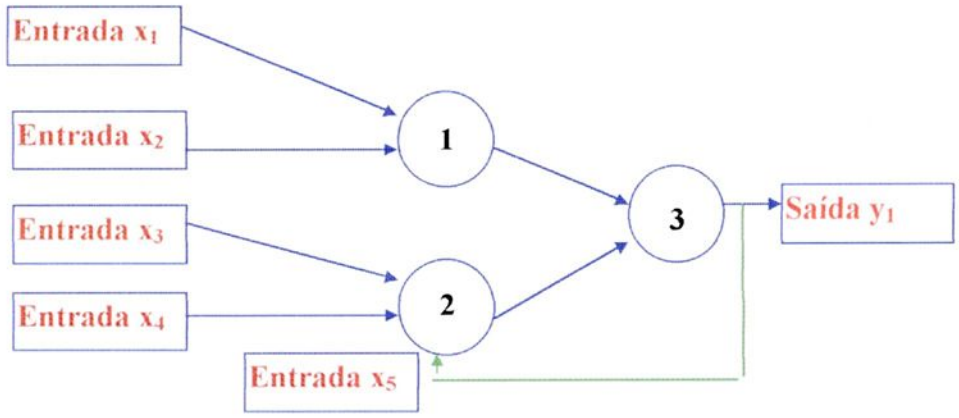

Figura 12.c. Múltiplas Camadas – Cíclica - Parcialmente Conectada, inspirada em Braga (1998, p. 11).

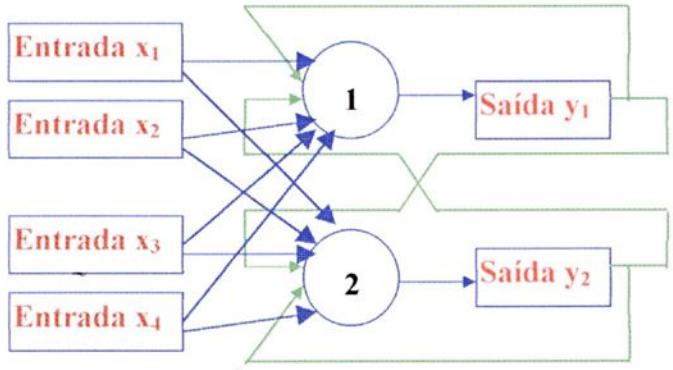

Figura 12.d. Camada Única – Cíclica - Totalmente Conectada, inspirada em Braga (1998, p. 11).

Apresentaram-se até o momento os fundamentos teóricos relativos às *Redes Neurais Artificiais*, passando-se pela *motivação* da continuidade dos estudos sobre redes neurais artificiais, pela composição do *neurônio artificial* e pelas *simulações de formação de padrões informacionais*, bem como pelo funcionamento de uma rede neural artificial, com as *principais arquiteturas* que a envolvem.

No próximo capítulo, serão apresentados alguns fundamentos de aprendizagem aplicada aos modelos supervisionados e não supervisionados de redes neurais artificiais.

cm 1 2 3 4 5 6 7 unesp 1 12 13 14 15 16 17 18

3

Aprendizagem

Segundo Kovács (1996), as redes neurais artificiais não são construídas com o propósito de resolver um problema específico, mas elas aprendem a resolvê-lo. Neste capítulo, se investigará os processos de aprendizagem, para que se tenha clareza acerca dos principais aspectos envolvidos nas redes neurais artificiais.

A primeira seção do presente capítulo apresenta os fundamentos teóricos das hipóteses sobre *aprendizagem*, bem como os dois fenômenos fundamentais dessa noção: o *refinamento de habilidades* e *aquisição de conhecimentos*.

O objeto da segunda seção, será uma breve discussão sobre os métodos de aprendizagem aplicados aos modelos de redes neurais artificiais de um modo geral. A seguir, serão apresentados os métodos de *aprendizagem supervisionada* e *não supervisionada*. O primeiro, se utiliza de um supervisor para alimentar a rede neural artificial; o segundo, faz uso do meio ambiente externo.

3.1. Fundamentos de Aprendizagem

Pretende-se aqui apresentar alguns fundamentos teóricos sobre aprendizagem, visto que as redes neurais artificiais fazem uso de métodos de

aprendizagem, por meio dos quais a rede recebe informações para gerar suas simulações de formação de padrões informacionais.

Soares (1997, p. 47) define aprendizagem como "a aquisição e a modificação de conceitos que a curto ou a longo prazo podem ou não modificar o comportamento dos indivíduos". Assim, de acordo com este autor, pode-se dizer que os comportamentos dos indivíduos são modificados ou não pela aquisição de novos conceitos ou mesmo pela modificação destes. É este o mecanismo que configura a aprendizagem em indivíduos.

Já em termos de sistemas, Rich e Knight (1993, p. 513) afirmam que um processo de aprendizagem é constituído por

> mudanças adaptáveis que ocorrem no sistema em sua execução, no sentido de que permitem que o sistema, na sua próxima vez, faça a mesma tarefa tirada do mesmo grupo, com mais eficiência e eficácia.

Segundo Rich e Knight (1993, pp. 513-515), verifica-se que, quanto mais um sistema biológico ou artificial realiza uma determinada tarefa, menos erros esse sistema efetuará ao realizar essa mesma tarefa. Isso se dá pelo fato de que o resultado de uma tarefa vai melhorando a cada nova realização e os erros ocorridos anteriormente vão sendo corrigidos. Mudanças adaptáveis vão ocorrendo, levando o sistema a melhorar mais e mais; o que não implica dizer que necessariamente um sistema, biológico ou artificial, melhora seu desempenho por realiza várias vezes a mesma tarefa. No exemplo dado sobre a criança e o fogo, bastou apenas uma aproximação do fogo, para a criança saber que aquilo não é bom. Em outros casos, mesmo praticando a mesma tarefa várias vezes, pode não

unesp

ocorrer uma melhora no desempenho, visto que podem existir características, tanto fisicas quanto mecânicas, que de certa forma impedem um melhor desempenho.

Ainda segundo Rich e Knight (1993, p. 514),

> a aprendizagem abrange uma ampla escala de fenômenos, na qual em uma das extremidades está o *refinamento de habilidades* e na outra extremidade está a *aquisição de conhecimentos*.

Na acepção de Rich e Knight (1993, p. 514), *refinamento de habilidades* refere-se à melhoria das habilidades que um sistema tanto biológico quanto artificial adquire por meio da prática. Ainda segundo os autores, quanto mais esses sistemas exercem suas tarefas, melhor será seu desempenho em determinadas tarefas. Tomando como exemplo o ser humano, pertencente ao sistema biológico, verifica-se que quanto mais exercer suas tarefas profissionais ou de lazer (digitar textos, andar de patins, jogar futebol, treinar malabarismo, etc), melhor será seu desempenho nessas tarefas, sem levar em consideração as limitações fisicas e/ou mentais. De outro lado, se novamente citarmos o exemplo da criança e do fogo, lembremo-nos de que bastou apenas uma vez para a criança saber que o fogo não faz bem.

Dependendo da atividade fisica exercida, uma pessoa com idade mais elevada poderá ter dificuldades em continuar melhorando seu desempenho, tendo já atingido seu máximo quando era dotado de boa forma fisica. Uma pessoa iniciante poderá melhorar a cada novo treinamento, podendo um dia atingir seu desempenho máximo.

Por outro lado, a *aquisição de conhecimentos*, tratada por Rich e Knight (1993, p. 514), é adquirida por meio da experiência, da vivência, mediante conselhos de outras pessoas, por memorização de conhecimentos, pela prática de tarefas e mediante sua própria experiência na solução de problemas. Ou seja, quanto mais se vive, mais conhecimentos se adquire.

Como exemplo prático, podem-se citar as tarefas profissionais das funções de programadores e analistas de sistemas, junto a um centro de computação de uma determinada empresa. Suponhamos que essa entidade já atua no mercado de computação há dez anos. Durante esses anos de atividades, tais profissionais adquiriram muita experiência, conhecimentos e habilidades, em virtude da prática constante e por meio de orientações de amigos e profissionais da área em questão. Hoje em dia, eles exercem suas tarefas com mais eficiência, resolvendo sozinhos a maioria de seus problemas, graças ao *refinamento de suas habilidades* e à *aquisição de conhecimentos*, adquiridos ambos durante os anos de prática constante.

Conforme argumentam Rich e Knight (1993, p. 513), "devemos ter em mente que a aprendizagem é em si mesma um processo de solução de problemas".

3.2. Aprendizagem nas Redes Neurais Artificiais

Segundo Haykin (2001), a propriedade mais importante para as redes neurais artificiais é a habilidade de aprender a partir de seu ambiente e melhorar o seu desempenho por meio da aprendizagem.

Haykin (2001, p. 75) define aprendizagem no contexto de redes neurais como

> um processo pelo qual parâmetros livres de uma rede neural são adaptados por meio de um processo de estimulação, pelo ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira na qual a modificação dos parâmetros ocorre.

Analisando a definição de Haykin, pode-se dizer que o processo de aprendizagem está envolvido no ambiente em que a rede está inserida. Esse ambiente acaba por estimular a rede, que sofre modificações em seus parâmetros. Esta, por sua vez, passa a responder de maneira nova ao ambiente que a estimulou, em razão das modificações ocorridas na sua estrutura interna.

Ainda, segundo Haykin (2001), um conjunto preestabelecido de regras para a solução de um problema de aprendizagem é denominado como um *algoritmo de aprendizagem*. Não há um algoritmo único para as redes neurais artificiais. Estes algoritmos diferem entre si pela maneira como é formulado o ajuste de um peso sináptico de uma unidade de processamento e pela maneira como a rede se relaciona com o seu ambiente.

Desse modo, nas redes neurais artificiais, a aprendizagem tem por objetivo determinar a intensidade das conexões entre as unidades de processamento, não sendo necessário fornecer o algoritmo correto para resolver um problema, uma vez que a rede o adapta por meio de um processo iterativo da intensificação das conexões entre as unidades de processamento. Essas unidades "guardam, ao final do processo, o conhecimento que a rede adquiriu do ambiente em que está operando" (Mendel & Mclaren, 1970, p. 287).

Disso se infere que esses algoritmos possuem instruções elaboradas e definidas pelo programador, não em termos de valores pré-determinados, mas em termos de ajustes de parâmetros e de funções matemáticas. Portanto, o programador não tem domínio exato sobre o resultado final após a execução desse tipo de algoritmo. Tais ajustes de parâmetros ocorrem por intermédio de informações relevantes, extraídas da rede neural artificial e dos padrões de informações apresentados a ela. As informações são fornecidas por um supervisor ou pelo meio ambiente externo, fato que leva a rede neural artificial a obter maior conhecimento do ambiente em que está operando.

Segundo Braga (1998, p. 15), "as redes neurais artificiais possuem a capacidade de aprender por exemplos e fazer interpolações e extrapolações do que aprendem". Então, quanto mais a rede pratica determinadas tarefas, com maior perfeição será capaz de realizá-las.

A capacidade de aprender das redes neurais artificiais, segundo Braga (1998, p. 15), resulta do

> processo pelo qual os parâmetros de uma rede neural artificial são ajustados por meio a uma forma continuada de estímulos pelo ambiente no qual a rede está operando, sendo o tipo específico de aprendizagem realizada definido pela maneira particular como ocorrem os ajustes realizados nos parâmetros.

Portanto, a rede adapta seus parâmetros conforme as informações provindas do meio ambiente no qual está inserida, promovendo-a a aprendizagem e a realização de determinadas tarefas. Por trás desse contexto está o algoritmo, mecanismo que realiza a adaptação da rede.

Conforme Braga (1998, p. 15), os algoritmos de aprendizagem

> são procedimentos e instruções bem elaborados, com o objetivo de adaptar os parâmetros de uma rede neural artificial, levando a mesma a aprender uma determinada tarefa, ou seja, são procedimentos finitários para resolver problemas.

Portanto, segundo o autor, é por intermédio dos algoritmos que as redes neurais artificiais aprendem uma determinada tarefa e elas próprias geram mudanças adaptáveis nesses algoritmos, visto que não são programadas para resolver um problema específico, mas treinadas para tal.

Faz-se necessário, neste momento, levantar algumas comparações a fim de se esclarecer a diferença entre os algoritmos utilizados para os treinamentos de redes neurais artificiais e os algoritmos utilizados em tarefas específicas, como as das máquinas de Turing.

Em ambos os casos os algoritmos são elaborados pelos programadores. A diferença é que, no primeiro caso (redes neurais artificiais), os algoritmos são elaborados utilizando-se procedimentos e funções matemáticas ajustáveis, isto é, procedimentos e funções que poderão sofrer ajustes e adaptações em seus parâmetros, por meio dos estímulos de entrada recebidos pela rede neural artificial e pelo supervisor, durante sua execução. Dessa forma, a rede neural artificial acaba por controlar a execução do algoritmo via estímulos de entrada recebidos, gerando mudanças adaptáveis no algoritmo. A rede neural artificial é, por esse motivo, considerada um sistema aberto. Deve-se lembrar, porém, que a rede não gera, não cria o algoritmo, apenas adapta o algoritmo inicialmente criado pelo programador de acordo com as informações recebidas pela rede.

cm 1 2 3 4 5 6 7 unesp 1 12 13 14 15 16 17 18

No segundo caso (máquinas de Turing), em relação às tarefas específicas, os algoritmos são elaborados com suas regras e instruções, pré-requisitos e valores pré-armazenados e pré-fixados, sendo tudo determinado pelo programador. Portanto, esse tipo de algoritmo não sofre influência dos estímulos por ele processados, isto é, os estímulos são somente processados, não tendo o poder de adaptar os procedimentos e funções dos algoritmos. Esses algoritmos não aprendem a resolver problemas em geral, eles apenas são elaborados para resolver determinados problemas, já que não possuem a flexibilidade de se auto-adaptarem, como ocorre nos algoritmos em redes neurais artificiais; portanto, são considerados sistemas fechados.

É oportuno acrescentar que, nas redes neurais artificiais, diversos algoritmos podem ser utilizados na aprendizagem, agrupando-se em dois métodos nos quais a aprendizagem possa ocorrer. Tratam-se do *método de aprendizagem supervisionada* e do *método de aprendizagem não supervisionada* ou *por auto-organização*.

3.2.1. Método de Aprendizagem Supervisionada

Na *aprendizagem supervisionada*, conhecida também por *aprendizagem com um professor*, é indispensável a presença de um supervisor para se garantir o funcionamento da rede.

Segundo Gonzales (1996, pp.279-280), um *método de aprendizagem supervisionada* é aquele em que um estímulo de entrada juntamente

com uma saída desejada são fornecidos à rede neural artificial por meio de um supervisor ou professor externo.

Segundo Haykin (2001), é o supervisor que tem o conhecimento do conjunto dos estímulos de entrada e de saída no ambiente em que a rede está inserida. Em virtude de seu conhecimento prévio, o supervisor é capaz de fornecer à rede uma resposta desejada àqueles estímulos de entrada. A resposta desejada representa a ação ótima, isto é, a resposta correta a ser realizada pela rede neural artificial.

Outro item importante, durante a fase de aprendizagem, é o *sinal de erro*. Segundo Haykin (2001, p. 88), o sinal de erro é definido como a "diferença entre a resposta desejada e a resposta real da rede". É de acordo com os estímulos de entrada e de saída combinados com o sinal de erro, que os parâmetros da rede são ajustados com o objetivo de fazer a rede emular o supervisor. Quando a rede alcançar a ação ótima de todos os estímulos de entrada, pode-se então dispensar o supervisor e deixar a rede neural lidar com o ambiente por si mesma. Essa forma de aprendizagem chama-se *aprendizagem por correção de erro*, por ser um sistema realimentado de laço fechado.

O supervisor, ou professor, é, portanto, aquele que indica explicitamente um comportamento bom ou ruim para a rede neural artificial. De acordo com os estímulos de entrada e saída fornecidos, o supervisor poderá direcionar o processo de aprendizagem, tornando-o mais fácil ou mais complexo, dependendo de quanto o estímulo de entrada está distanciado da saída, e também

unesp

do número de camadas intermediárias existentes entre a camada de entrada e a camada de saída.

Baseando-se nas definições anteriores, pode-se dizer que o objetivo do método de aprendizagem supervisionada é ajustar os parâmetros da rede neural artificial, de forma a encontrar uma ligação entre os estímulos de entrada e de saída fornecidos à rede neural artificial. A figura a seguir ilustra uma rede neural artificial de aprendizagem supervisionada.

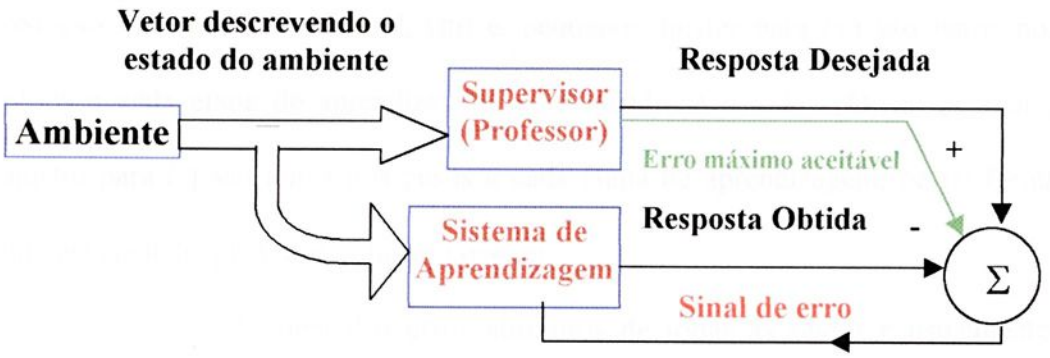

Figura 13. Exemplo de uma rede neural artificial com método de aprendizagem supervisionada (Haykin, 2001, p. 88).

Analisando a Figura 13, verifica-se que a rede neural artificial obtém resposta do estado do ambiente e dos estímulos de entrada processados. Compara-se, então, esta resposta à resposta desejada, via função somatória, encaminhada à rede neural artificial pelo supervisor, que também forneceu dados sobre o erro máximo que a rede neural artificial deverá considerar. O erro, resultante da comparação da resposta obtida à resposta desejada, deverá ser menor ou igual ao erro máximo aceitável, fornecido à rede neural artificial pelo supervisor, para que a rede neural artificial suspenda o processo de aprendizagem.

Em caso contrário, o erro é enviado à rede neural artificial que ficará em execução ou treinamento, até que o erro fique menor que o valor do erro máximo aceitável.

Em cada estímulo de entrada, submetido à rede neural artificial, compara-se a resposta desejada à resposta obtida e ajustam-se os pesos das conexões para reduzir o erro. Os pesos são aproximados para (+), caso a resposta obtida não seja igual à resposta desejada e seja menor que zero (0); e para (-), caso a resposta obtida não seja igual à resposta desejada e seja maior que um (1). A redução do erro é incremental, isto é, pequenos ajustes para (+) são feitos nos pesos a cada etapa de aprendizagem, adicionando-se o valor (1); ou pequenos ajustes para (-) são feitos nos pesos a cada etapa de aprendizagem, de tal forma que ela caminhe para uma solução possível.

A soma dos erros absolutos de todas as saídas é usualmente, utilizada como medida de desempenho da rede neural artificial e considerada como função de custo a ser reduzido pelo algoritmo de aprendizagem. Esse método de aprendizagem necessita, portanto, da existência de um supervisor para que a rede neural artificial possa aprender novas estratégias para lidar com situações desconhecidas.

Como exemplos de algoritmos de métodos de aprendizagem supervisionada, serão abordados no capítulo 4, seção 4.1, dois modelos de redes neurais artificiais: o modelo *Perceptron de Camada Única* e o modelo *Perceptron de Múltiplas Camadas*. Esse último utiliza-se do algoritmo de aprendizagem Retropropagação (*Backpropagation)*.

3.2.2. Método de Aprendizagem sem um Supervisor ou por Auto-organização

Como o próprio nome já diz, esse método difere-se do anterior uma vez que ele não possui um supervisor para acompanhar o processo de aprendizagem. A rede neural artificial não sofre influência de um supervisor, mas recebe estímulos de entrada provindos do meio ambiente externo. Trata-se de método conhecido também como aprendizagem por *Auto-organização*.

Segundo Haykin (2001), o método de *aprendizagem sem um supervisor* pode ser dividido em dois outros: *aprendizagem por reforço* e *aprendizagem não supervisionada*.

- **Aprendizagem por reforço**

Conforme Haykin (2001), nesse método, a aprendizagem de um mapeamento de estímulos de entrada e de saída é realizada por meio da interação contínua com o ambiente, visando minimizar um índice escalar de desempenho. Ainda segundo Haykin, esse tipo de aprendizagem é construído em torno de um *crítico* que converte um *sinal de reforço primário* recebido do ambiente em um sinal de reforço de melhor qualidade, denominado *sinal de reforço heurístico*.

Na realidade, o papel do *crítico* é observar uma seqüência temporal de estímulos de entrada recebidos do ambiente, sendo que a cada passo dessa seqüência, ele possa observar se a nova resposta tem uma melhor qualidade do que a resposta do passo anterior, e assim sucessivamente, até obter-se uma resposta com mais qualidade. Infere-se de que ai ele é o supervisor.

A Figura 14 ilustra esse tipo de aprendizagem.

unesp

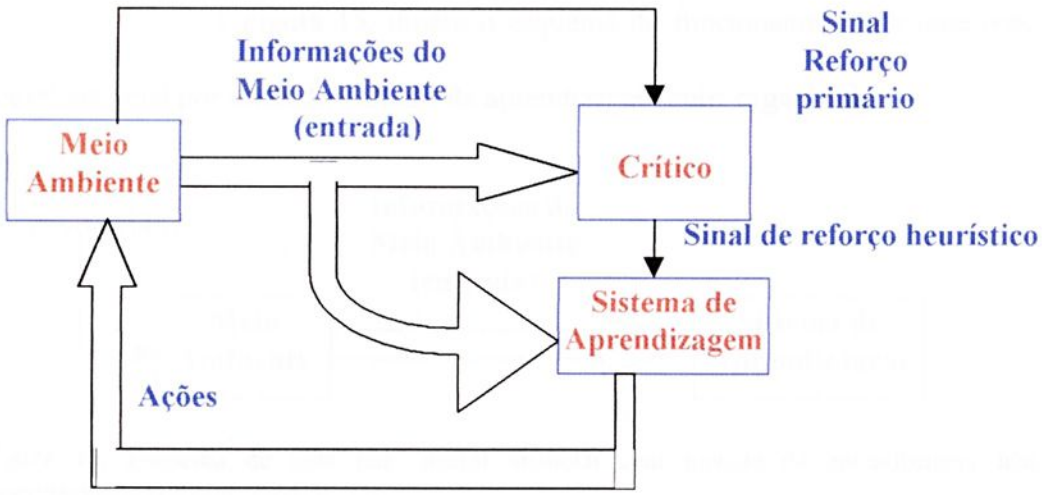

Figura 14. Esquema de uma rede neural artificial com método de aprendizagem por reforço (Haykin, 2001, p. 90).

- **Aprendizagem não supervisionada**

Segundo Haykin (2001), nesse método, não há um supervisor externo. Em vez disso, são dadas condições para a realização de uma medida independente da tarefa, na qualidade da representação que a rede deve aprender. Os parâmetros livres da rede são otimizados em relação a esta medida. Conforme Haykin ainda, uma vez que a rede tenha se ajustado às regularidades estatísticas dos dados de entrada, ela desenvolve a habilidade de formar representações internas para codificar as características da entrada e, desse modo, criar automaticamente novas classes.

Segundo Gonzales (1996, p. 280), no método de aprendizagem por auto-organização,

> nenhum estímulo específico de entrada ou saída é fornecido de antemão à rede, que terá que classificar, criar ou detectar um estímulo, por meio do reconhecimento de regularidade na informação provinda do meio ambiente, sem a ajuda de um supervisor.

A Figura 15, ilustra o esquema do funcionamento de uma rede neural artificial por meio do método de aprendizagem auto-organizado.

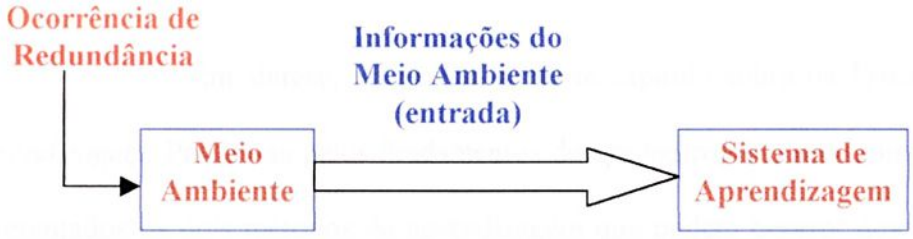

Figura 15. Esquema de uma rede neural artificial com método de aprendizagem não supervisionada (Haykin, 2001, p. 91).

Analisando a Figura 15, verifica-se que um modelo de rede neural artificial de aprendizagem não supervisionada possui ocorrência de redundância, isto é, as informações provindas do meio ambiente externo são introduzidas na rede neural artificial inúmeras vezes. Essas informações são processadas, obtendo-se uma resposta que muitas vezes se aproxima da resposta desejada, mas que em outras vezes, pode ficar distante desta.

A partir do momento em que a rede neural artificial estabelece uma comunicação extraindo informações disponíveis no meio ambiente externo, vindas com os estímulos de entrada, desenvolve-se nela uma habilidade de formar representações internas (padrões informacionais) para codificar características dos estímulos de entrada e criar novas classes ou grupos automaticamente. Esse tipo de aprendizagem só se torna possível quando existe redundância dos estímulos de entrada. Sem redundância, seria muito difícil - não impossível - encontrar quaisquer padrões ou características dos estímulos de entrada.

Exemplos de algoritmos (o modelo *SOFM* de Kohonen) de métodos de aprendizagem não supervisionada serão vistos no capítulo 4, seção 4.2.3.

Em síntese, investigou-se neste capítulo sobre os *Processos de Aprendizagem*. Passou-se pelos fundamentos de *aprendizagem*, bem como, foram apresentados os dois métodos de aprendizagem que podem ocorrer nos modelos de redes neurais artificiais: a *aprendizagem supervisionada* e a *aprendizagem não supervisionada*. Ainda foram abordadas as características dos modelos de redes neurais artificiais supervisionados e dos modelos de redes neurais não supervisionados ou auto-organizados.

A seguir, apresentaremos modelos de redes neurais artificiais com aprendizagem supervisionada e não supervisionada.

4

Modelos de Redes Neurais Artificiais

A primeira seção deste capítulo apresenta dois modelos de redes neurais artificiais com *aprendizagem supervisionada*; trata-se do *Perceptron de Camada Única* e do *Perceptron de Múltiplas Camadas*.

Um modelo de rede neural artificial com *aprendizagem não supervisionada* é o tema da segunda seção; trata-se do modelo *SOFM* de Kohonen.

A terceira seção mostrará um treinamento de uma rede neural artificial com aprendizagem supervisionada para reconhecimento e classificação de padrões informacionais.

Um outro treinamento via aprendizagem não supervisionada para detecção ou, até mesmo criação e classificação de padrões informacionais é o assunto da Quarta e última seção do capítulo.

4.1. Modelos com Aprendizagem Supervisionada

Serão abordados aqui os modelos *perceptron de camada única* e *perceptron de múltiplas camadas* (algoritmo *backpropagation*).

4.1.1. *Perceptron de Camada Única*

O modelo *perceptron de camada única*, originalmente proposto por Frank Rosenblatt em 1957, apresenta uma arquitetura de padrões

> composta por uma estrutura de rede neural artificial tendo unidades básicas de processamento (MCP) e uma regra de aprendizagem, na qual o objetivo é aprender a classificar estímulos de entrada por meio de um treinamento supervisionado (Braga, 1998, pp. 32-33).

Segundo Marar (1997, p. 73), o *perceptron de camada única* é o exemplo mais simples de uma rede *feedforward*, o qual consiste de uma única unidade de processamento com *n* estímulos de entrada e um único estímulo de saída.

Conforme Haykin (2001, p. 143), o *perceptron de camada única* "é a forma mais simples de uma rede neural artificial, usada para a classificação de padrões ditos linearmente separáveis, isto é, padrões que se encontram em lados opostos de um hiperplano".

Segundo esse mesmo autor, o *perceptron de camada única* compõe-se basicamente de uma única unidade de processamento com pesos sinápticos ajustáveis e de um valor de inclinação.

A Figura 16 e a função correspondente ilustram um modelo *perceptron de camada única*.

Analisando-se a Figura 16, pode-se observar que o modelo *perceptron* exemplificado é composto por dois estímulos de entrada, uma inclinação inicial fixada, terminais de entrada, pesos iniciais, pesos finais, uma unidade de processamento, uma função soma e um terminal de saída.

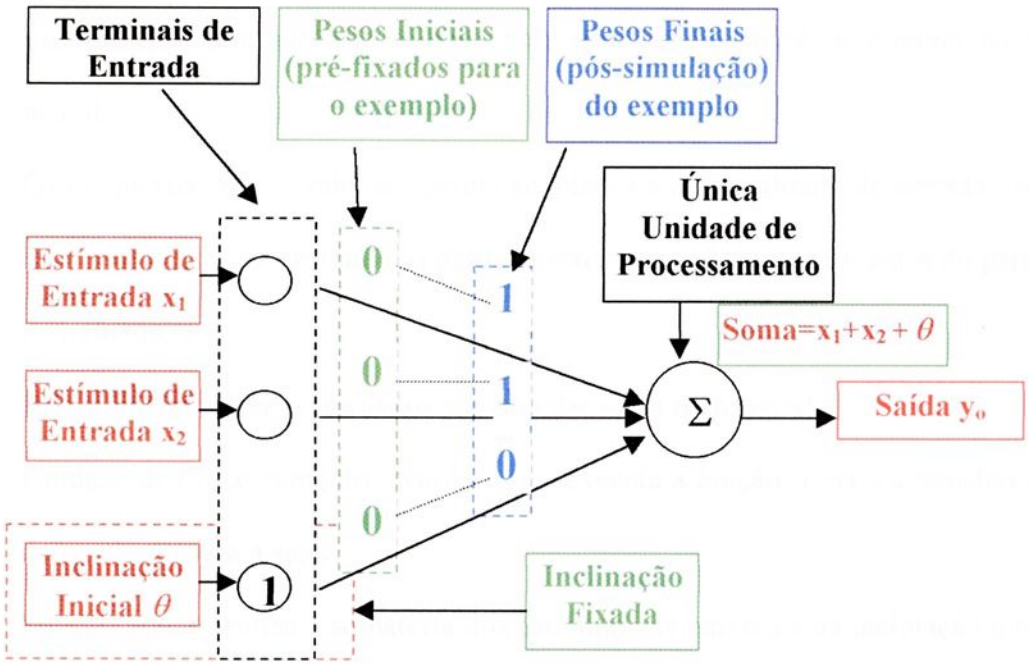

Figura 16. Modelo de uma rede neural artificial *Perceptron de Camada Única*, inspirada de Braga (1998, p. 28).

$$y = \sum_{i=1}^{n} WiXi + \theta$$

A seguir, é apresentada a função de cada componente utilizado no modelo *perceptron de camada única*:

Estímulos de entrada: são os sinais, os dados, que entram na rede neural artificial. No exemplo são definidos como x_1, x_2.

Inclinação Inicial (bias): tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo do fato de ela ser positiva ou negativa, respectivamente. Nesse exemplo, definimos a inclinação inicial com valor um (1), antes do período de treinamento. No exemplo é definida como θ.

Terminais de Entrada: são o meio pelo qual os estímulos se inserem na rede neural.

Pesos Iniciais: são os valores a serem acoplados a cada estímulo de entrada. Nesse exemplo foram usados todos os pesos iniciais com valor zero (0), antes do período de treinamento.

Pesos Finais: valores retornados pós-simulação da rede neural.

Unidade de Processamento: neurônio que executa a função soma e a transferência do resultado para a saída.

Função Soma: realiza a somatória dos estímulos de entrada e da inclinação inicial, multiplicados pelos seus respectivos pesos.

Terminal de Saída: resposta obtida da simulação, decorrente da somatória dos estímulos de entrada e da inclinação inicial multiplicadas pelos seus respectivos pesos. No exemplo é definida como saída y_0.

Analisando a função soma, o valor resultante é aplicado (comparado) ao limiar fixado na função de ativação utilizada pelo modelo. A unidade de processamento produzirá uma saída igual a +1 se a entrada do limiar for positiva e –1 se ela for negativa. Por meio da função soma, obtém-se o valor resultante que será aplicado à função de ativação. O objetivo desse modelo é, pois, classificar os estímulos de entrada fornecidos à rede em duas classes: classe C_1 para a saída y com valor +1 e classe C_2 para o y com valor –1.

Minsky e Papert (1969) argumentaram enfaticamente quanto às limitações básicas do *perceptron de camada única*, porque somente resolvia problemas linearmente separáveis, como por exemplo os operadores lógicos AND

e OR, demonstrando que as redes *perceptron de múltiplas camadas* eram capazes de melhor performance e de proporcionar melhor solução para problemas não linearmente separáveis, como por exemplo o operador lógico XOR. Contudo, naquela época, o algoritmo *backpropagation* de aprendizagem para as redes *perceptron de múltiplas camadas* ainda era desconhecido.

Na subseção seguinte, será apresentado o algoritmo *backpropagation* aplicado às redes *perceptron de múltiplas camadas*.

4.1.2. Algoritmo *Backpropagation* e *Perceptron de Múltiplas Camadas*

Conforme relata Marar (1997), a utilização do algoritmo *backpropagation* teve seu início em 1974 com a tese de doutorado de Paul Werbos, mas tornou-se popular com os trabalhos de Rumelhart, Hinton e Willians.

Em 1986, o livro PDP (Parallel Distributed Processing), de Rumelhart, Hinton e Willians, mostrou que as redes *perceptron de múltiplas camadas, feedforward* e totalmente conectadas podiam ser amplamente utilizadas, graças ao algoritmo de retropropagação de erro (*error backpropagation)*, o qual consta de dois passos: um passo para frente (*propagação*) e um para trás (*retropropagação*). Esse algoritmo é baseado na *aprendizagem por correção de erro* (ver Capítulo 3, subseção 3.2.1).

Analisando os dois passos desse modelo, pode-se afirmar que no *passo para frente* um padrão de entrada, e os pesos fixos são fornecidos à rede e seu efeito se propaga camada por camada. No *passo para trás*, todos os pesos

sinápticos são ajustados de acordo com uma regra de correção de erro. Especificamente a resposta obtida é subtraída da resposta desejada a fim de produzir um *sinal de erro*.

Segundo Kovács (1996), o algoritmo *backpropagation* é aplicado às redes neurais artificiais de múltiplas camadas do tipo *perceptron*. Esse tipo de rede possui mais de uma camada de unidades de processamento, conhecida por *Perceptron de Múltiplas Camadas* (MLP).

Segundo Haykin (2001), a rede *perceptron de múltiplas camadas* forma-se de um conjunto de unidades de processamento sensoriais que constituem a camada de entrada, uma ou mais camadas intermediárias e uma camada de saída. Estes modelos de *perceptron* representam uma generalização do *perceptron de camada única.*

Conforme considerações de Marar (1997) e Kovács (1996), esse algoritmo opera numa seqüência de cinco passos durante o treinamento de uma rede neural artificial de múltiplas camadas do tipo *perceptron*:

- **Primeiro passo:** corresponde a um estímulo de entrada que é apresentado à rede neural artificial;
- **Segundo passo:** os fluxos das atividades resultantes fluem por meio da rede neural artificial de camada em camada, até que a resposta da rede seja gerada;
- **Terceiro passo:** compara-se a resposta obtida pela rede neural artificial à resposta desejada para aquele estímulo de entrada particular. Se a resposta obtida não for igual à resposta desejada, um erro é gerado e propagado de

volta à rede. Caso contrário, a rede prossegue o treinamento, passando esse resultado à unidade de processamento posterior;

- **Quarto passo:** o erro é propagado por intermédio da rede neural artificial, vindo a partir da camada de saída até a camada de entrada, com os pesos das conexões das unidades de processamento sendo modificados com o erro retropropagado;
- **Quinto passo:** repetem-se os passos 1 ao 4 sobre o conjunto de treinamento, até que o erro global seja aceitável.

A Figura 17, de própria autoria, ilustra o esquema de funcionamento de uma rede neural artificial *perceptron de múltiplas camadas*, por meio do algoritmo *backpropagation*, de acordo com os cinco passos mencionados anteriormente.

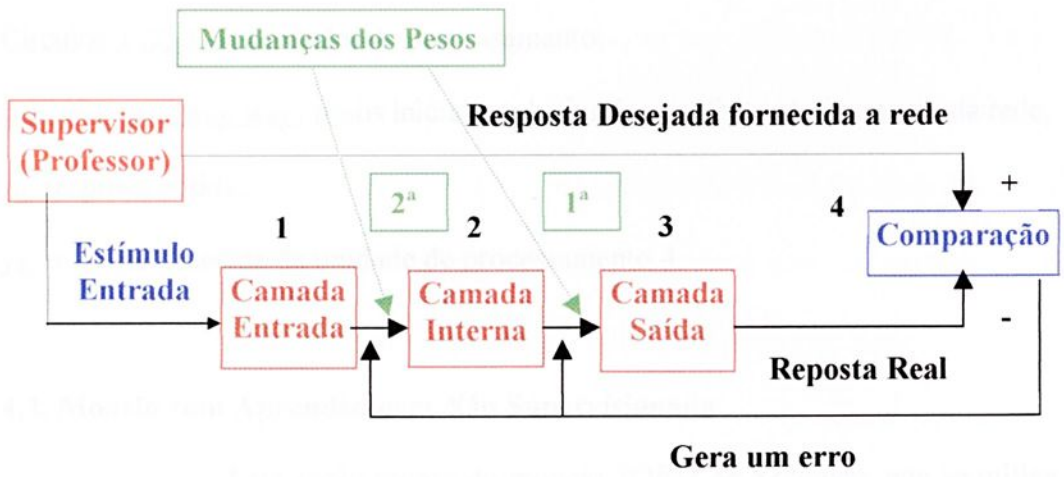

Figura 17. Esquema de uma rede neural artificial *Perceptron de Múltiplas Camadas*.

Segundo as definições anteriores, pode-se classificar os passos 1, 2 e 3 como fase de uso da rede e os passos 1 ao 4.como fase de aprendizagem

Na Figura 18, de própria autoria, mostra-se o plano anterior em um esquema prático de uma rede neural artificial *perceptron de múltiplas camadas* com algoritmo de aprendizagem *backpropagation*.

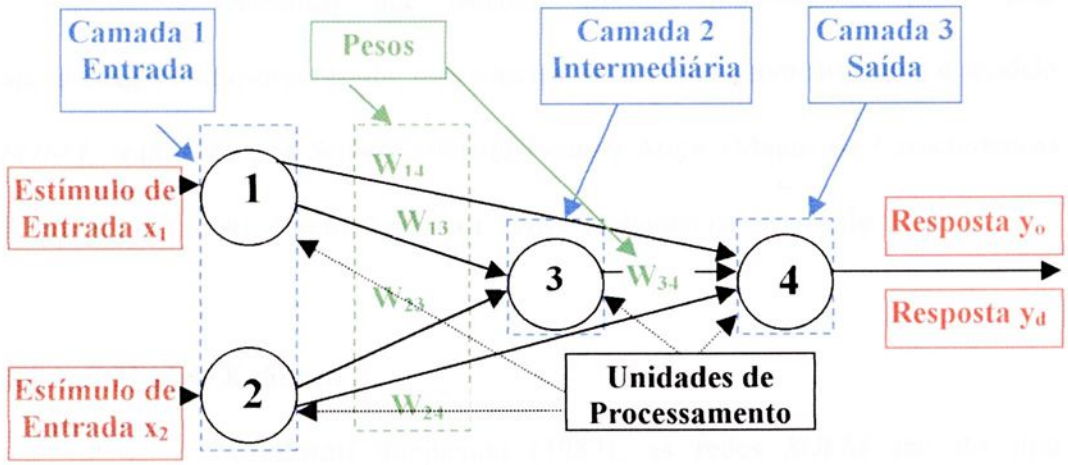

Figura 18. Modelo de uma rede neural artificial *Perceptron de Múltiplas Camadas*.

x_1 e x_2 : estímulos de entrada;

Círculos 1 ,2, 3, 4: unidades de processamento;

w_{31}, w_{32}, w_{41}, w_{42}, w_{43} : pesos iniciais, substituídos conforme treinamento da rede;

y_o: resposta obtida;

y_d: resposta desejada da unidade de processamento 4.

4.2. Modelo com Aprendizagem Não Supervisionada

Esta seção tratará do modelo *SOFM* de Kohonen, que se utiliza da aprendizagem não supervisionada ou auto-organizada. O modelo *SOFM* reconhece e classifica padrões informacionais sem a interferência de um supervisor, ao contrário dos modelos supervisionados, nos quais existe a intervenção de um supervisor.

Conforme Haykin (2001), o principal objetivo do modelo *SOFM* é produzir um mapa topologicamente ordenado que localize otimamente estímulos de entrada (vetor), facilitando a compreensão dos dados.

Sabemos que existem diversos modelos de redes com aprendizagem auto-organizada, mas este trabalho versará somente sobre o modelo *SOFM*, conhecido por *Self-Organising Feature Maps* (Mapas de Características Auto-Organizados), desenvolvido por Teuvo Kohonen na década de 80.

4.2.1. *SOFM* de *Kohonen*

Conforme Lippmann (1987), as redes *SOFM* são do tipo *feedforward* (acíclica – alimentação adiante) (ver Capítulo 2, seção 2.5) com aprendizagem não supervisionada. Essas redes usam simples unidades de processamento adaptativas para receber sinais de um ambiente externo. Esses sinais compõem-se de medidas ou dados, como freqüência ou situação.

Baseado em Haykin (2001), pode-se afirmar que esse modelo é considerado uma classe especial de *mapas neurais*, baseada na *aprendizagem competitiva*. Nela, as unidades de processamento da camada de saída competem entre si para serem ativadas ou disparadas, por meio de uma *função de ativação*. O resultado é que apenas uma unidade de processamento de saída sairá *vencedora*. Essa competição recebe o nome de *"o vencedor leva tudo"*. Trata-se de um tipo de competição induzido por conexões laterais inibitórias (negativas) entre as unidades de processamento.

Já Kohonen (1987) observa que o esquema básico de seu modelo *SOFM* faz com que unidades de processamento da camada intermediária disputem entre si a representação de informação fornecida pelas unidades de processamento de entrada, a fim de se tornarem a *unidade vencedora*. Em havendo uma unidade de processamento vencedora, esta é reajustada para responder ainda melhor ao estímulo recebido. Este modelo não somente ajusta o vencedor, mas também as unidades de processamento vizinhas à unidade vencedora.

Segundo Lippmann (1987), o modelo *SOFM* é baseado no mapa topológico presente no córtex cerebral, que mantém uma ordem entre suas unidades de processamento (neurônios). Sabe-se que o cérebro humano possui áreas responsáveis por funções específicas: áreas dedicadas à fala, à visão, ao controle motor, ao tato, etc. Essas áreas possuem sub-áreas que mapeiam, de forma ordenada, as entradas do orgão sensorial representado por ela.

Apresentadas as definições dos módulos, pode-se observar agora como cada unidade de processamento influencia o estado de ativação de suas unidades de processamento vizinhas. De acordo com Braga (1997), as três formas são: forma *excitatória*; *inibitória* e *levemente excitatória*.

- **Forma Excitatória:** quando as unidades de processamento vizinhas estão bem próximas à unidade de processamento que será a futura vencedora, ou seja, dentro de uma área de raio *R1*.
- **Forma Inibitória:** quando as unidades de processamento vizinhas estão fora da área anterior (*R1*), mas dentro de uma segunda área de raio (*R2*), sendo $R2 > R1$.

unesp

- **Forma Levemente Excitatória:** quando as unidades de processamento vizinhas estão fora das áreas anteriores (*R1* e *R2*), mas dentro de uma área de raio *R3*, sendo *R3* > *R2*.

A Figura a seguir, conhecida por *chapéu mexicano*, ilustra o processo de ativação definido anteriormente.

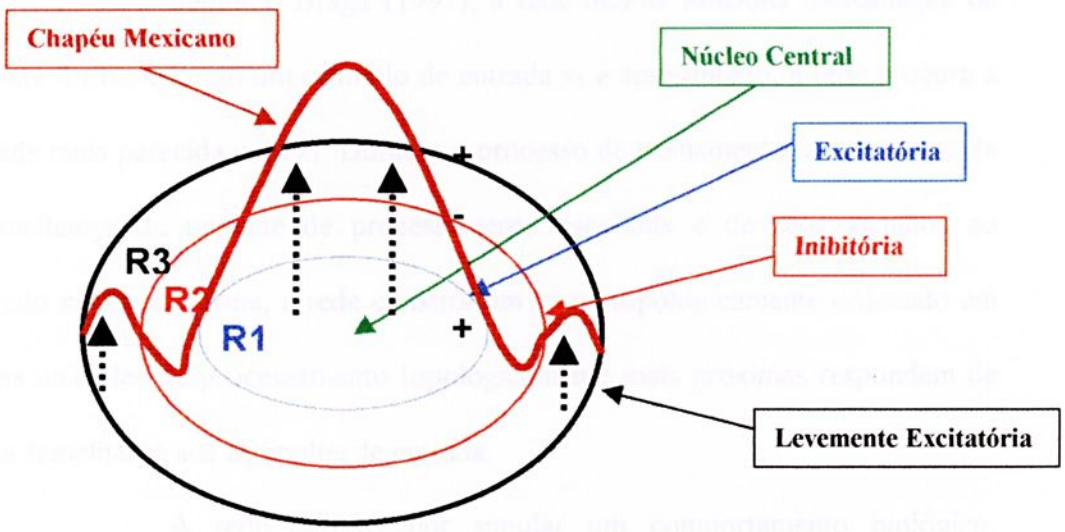

Figura 19. Tipos de estímulos produzidos de acordo com a função chapéu mexicano, extraída de Braga (1998, p. 105).

A Figura 19 indica que, estando as unidades de processamento mais próximas do núcleo central da função, cuja curva apresenta a forma de um chapéu mexicano, maiores serão os seus estados de ativação. Sendo assim, as unidades de processamento (neurônios) que estiverem dentro do raio *R1* possuem um estado de ativação *excitatório*; as que estiverem dentro do raio *R2* possuem um estado de ativação *inibitório* e as que estiverem dentro do raio *R3* possuem um estado de ativação *levemente excitatória*.

Portanto, quanto mais próxima do núcleo central, mais rapidamente a unidade de processamento sairá vitoriosa perante aquelas que estão concorrendo com ela. Esse processo de disputa de ativação recebe o nome de *aprendizagem competitiva*.

4.2.2. Mecanismo da rede *SOFM*

Segundo Braga (1997), a rede *SOFM* funciona basicamente da seguinte forma: quando um estímulo de entrada x_1 é apresentado, a rede procura a unidade mais parecida com x_1. Durante o processo de treinamento, a rede aumenta a semelhança da unidade de processamento escolhida e de seus vizinhos ao estímulo x_1. Dessa forma, a rede constrói um mapa topologicamente ordenado em que as unidades de processamento topologicamente mais próximas respondem de forma semelhante aos estímulos de entrada.

A rede *SOFM*, por simular um comportamento biológico, apresenta viabilidade biológica. Assim como no córtex cerebral, unidades de processamento ativas apresentam conexões mais fortes com unidades de processamento que estão fisicamente mais próximas.

Segundo Kohonen (1989), o mecanismo da rede *SOFM* funciona da seguinte forma: os *pesos sinápticos* iniciam contendo valores aleatoriamente baixos; um *estímulo de entrada* x é provido para a rede, sem que se especifique a saída desejada. Segundo ainda esse autor, o fato de não se especificar uma saída desejada caracteriza a rede como não supervisionada ou auto-organizada. O estímulo de entrada x é descrito como um vetor $\vec{x}=(x_1, x_2, x_3, ..., x_n)$. Quando o

estímulo de entrada é introduzido na rede, ele é disputado pelas unidades de processamento de saída *y*. A unidade de saída que melhor responder ao estímulo de entrada, será definida como unidade de saída vencedora.

A rede é considerada treinada depois que todo o conjunto de estímulos de entrada do vetor $\vec{x}$ de treinamento tiver sido apresentado à rede e satisfeitos os critérios de aprendizagem.

4.2.3. Algoritmo da rede *SOFM*

Conforme Lippmann (1987), o algoritmo de aprendizagem da rede *SOFM* de Kohonen também definido como algoritmo de *aprendizado competitivo*, opera numa seqüência de seis passos durante o treinamento, os quais podem ser assim descritos:

- **Primeiro passo:** corresponde à iniciação dos pesos das conexões $\vec{w}=(x_{11}, x_{21},..., x_{n1})$, com valores aleatórios próximos de zero. Os pesos aleatórios, em certos casos, levam a um processo muito lento de convergência. Dessa forma, pode ocorrer a iniciação da rede pela distribuição dos pesos uniformemente pelo espaço de pesos para acelerar o processo. Também no primeiro passo é iniciado o raio da vizinhança da unidade de processamento *j*, chamado de $Vj(t)$, para um valor de raio abrangente.
- **Segundo passo:** realiza-se a apresentação dos estímulos de entrada $\vec{x}=(x_1(t), x_2(t),..., x_n(t))$, sendo $\vec{x}$ o vetor das unidades de processamento que estão no tempo *t*.

unesp

- **Terceiro passo:** efetua-se o cálculo da distância euclidiana entre o padrão do estímulo de entrada e cada unidade de processamento da rede. Ou seja, toma-se uma unidade de processamento com índice 2, os estímulos de entrada apresentados por X=[1,3,2,4] e os respectivos pesos, apresentados por W=[0,1,0,1]. Segundo Haykin (2000, p. 51), a *distância euclidiana* entre o par de vetores, padrão do estímulo de entrada e pesos é definida por

$$d(X_i, W_j) = ||X_i - W_j||$$

$$= \left[\sum_{k=1}^{n} (X_{ik} - W_{jk})^2 \right]^{1/2}$$

- **Quarto passo:** seleciona-se a unidade de processamento que apresentou a saída com menor distância euclidiana. Esta unidade de processamento é designada como a *unidade vencedora*.
- **Quinto passo:** atualizam-se os pesos da unidade de processamento juntamente com os pesos de todas as unidades de processamento que estão dentro da vizinhança.
- **Sexto passo:** se o número de iterações realizadas até o momento atingiu o número especificado no início, encerra-se a fase do treinamento. Caso contrário, repete-se o processo a partir do segundo passo.

Analisando a figura seguinte, observa-se que a unidade de processamento do centro dela, definida com a cor preta, é a *unidade vencedora*. As demais unidades, definidas com a cor vermelha, localizadas bem próximas à

unidade vencedora, são as unidades vizinhas, ou seja, a vizinhança da unidade vencedora. Na parte inferior da figura tem-se a apresentação dos estímulos de entrada da rede, codificados pelo vetor de entrada $x = (x_1(t), x_2(t),..., x_n(t))$.

A Figura 20 ilustra a arquitetura de uma típica rede *SOFM*.

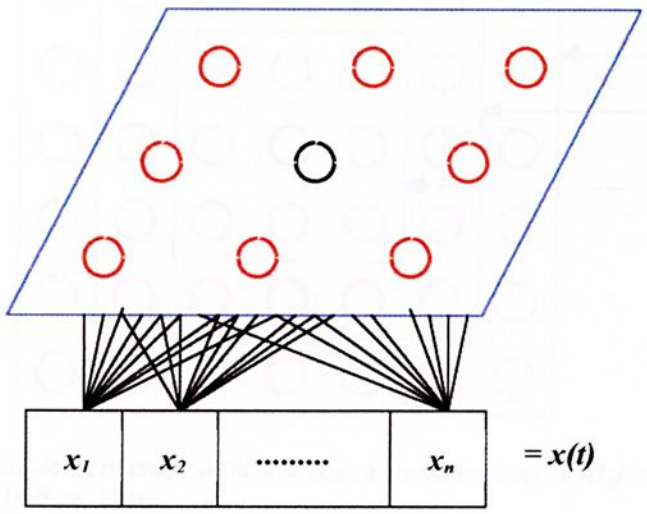

Figura 20. Exemplo de uma rede *SOFM* de Kohonen típica (bidimensional), extraída de Braga (1998, p. 106).

Segundo Kohonen (1987), a arquitetura da rede *SOFM* é determinada pelo modo como são definidos os seus vizinhos. No início, a vizinhança é extensa e não muito bem definida. A importância de a vizinhança ser inicialmente extensa propicia melhor treinamento da rede, já que ao se utilizar de um número maior de unidades vizinhas ela se ordenará de maneira global. Segundo Kohonen (1987), a melhor auto-organização ocorre quando a vizinhança começa extensa e diminui com o tempo variável. Se a vizinhança se inicia muito pequena em relação ao mapa, este não se ordenará globalmente.

A Figura 21 ilustra uma simulação do processo de vizinhança global, que se inicia no tempo $Vj(t_0)$ e decresce formando $Vj(t_1)$, chegando ao

tempo $Vj(t_2)$. As unidades pertencentes a esse tempo final são definidas como unidades vizinhas.

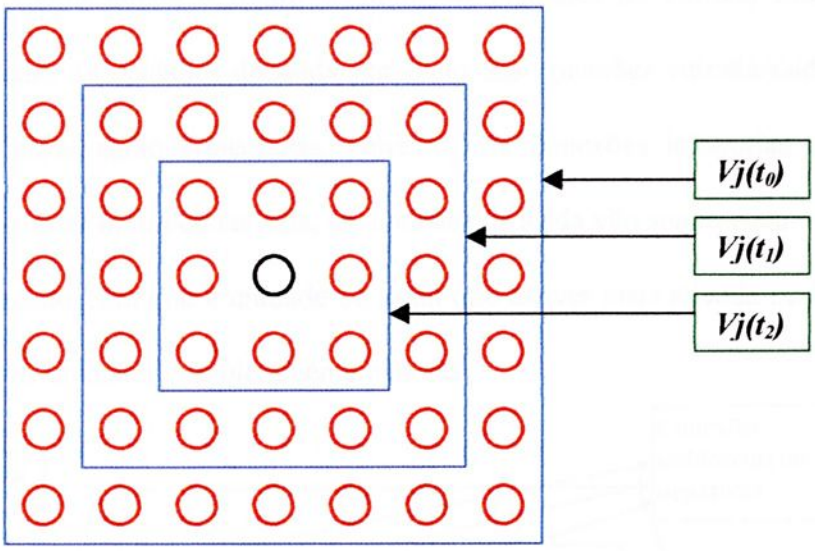

Figura 21. Topologia de vizinhança, de acordo com o tempo variável, $Vj(t_0) > Vj(t_1) > Vj(t_2)$, extraída de Braga (1998, p. 110).

A regra de atualização dos pesos no quinto passo faz com que a unidade de saída fique propensa a disparar quando perceber a mesma entrada novamente. Se a mesma entrada for apresentada muitas vezes, a unidade de saída ajustará seus pesos para a ativação máxima daquela entrada. Porém, como os vetores de entrada chegam misturados, as unidades de saída nunca se estabelecem em um conjunto perfeito de pesos. A expectativa é que cada uma encontre um grupo natural de vetores de entrada e gravite em sua direção, ou seja, na direção das ativações fornecidas, quando aquelas entradas lhes forem apresentadas. O algoritmo pára quando as mudanças de pesos ficam muito pequenas.

As unidades de processamento da camada de saída competem, portanto, entre si para se tornarem ativas, isto é, para ver qual delas gera o maior

valor de saída. Assim, para cada padrão de entrada apenas uma unidade de processamento de saída, por grupo, se torna ativa.

Como ilustra a figura seguinte, as unidades de entrada estão diretamente atreladas às unidades de saída por meio das conexões entrada/saída. As unidades de saída estão conectadas, entre si, via conexões inibitórias ou negativas. Ao longo das linhas de entrada, as unidades de saída vão sendo cada vez mais ativadas, fazendo com que a unidade de saída que estiver mais ativada acabe por deixar as outras unidades completamente inativas.

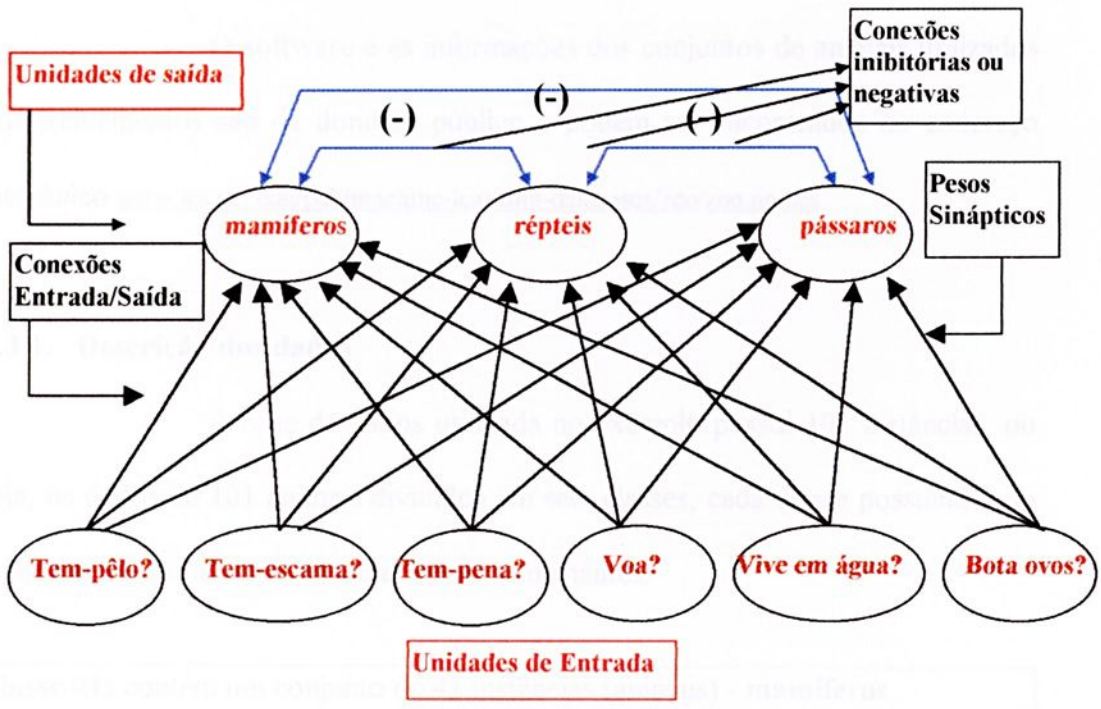

Figura 22. Exemplo de uma rede de aprendizado competitivo, inspirada em Gonzales (1996).

4.3. Treinamento de uma Rede Neural Artificial *Backpropagation* para reconhecimento e classificação de padrões informacionais

As redes neurais artificiais são consideradas ferramentas de grande capacidade de aprendizagem, via treinamento, em diversas áreas do conhecimento. Elas são desenvolvidas para reconhecer e classificar padrões informacionais. Apresenta-se aqui uma rede cujo objetivo é reconhecer e classificar padrões informacionais de animais, dentro de uma classe deles, por meio de suas características físicas e biológicas; segundo as características dos animais (ter penas, botar ovos, ter patas, respirar) é que se definirá a que classe cada espécie pertence.

O software e as informações dos conjuntos de animais utilizados nos treinamentos são de domínio público e podem ser encontrados no endereço eletrônico www.ics.uci.edu/pub/machine-learning-databases/zoo/zoo.names.

4.3.1. Descrição dos dados

A base de dados utilizada no exemplo possui 101 instâncias, ou seja, os dados de 101 animais divididos em sete classes, cada classe possuindo um conjunto de animais com características semelhantes.

Classe 01: contém um conjunto de 41 instâncias (animais) - **mamíferos**
Conjunto de animais: porco da terra (mamífero africano), antílope, urso, javali, búfalo, bezerro, capivara, chita (leopardo de caça da Índia), cervo, golfinho, elefante, fruitbat, girafa, garota, cabra, gorila, *hamster* (rato da Índia), lebre, leopardo, leão, *lince* (gato selvagem), marta, toupeira, mangusto (fuinha, animal presa), gambá, *órix*, ornitorrinco, esquilo, pônei, marsopa, onça parda, bichano, *raccoon* (animal carnívoro, família do urso), rena, foca, leão marinho, esquilo, vampiro, arganaz (rato silvestre), *wallaby* (pequeno canguru), lobo.

Classe 02: contém um conjunto de 20 instâncias (animais) - **aves**
Conjunto de animais: galinha (frango), corvo, pombo, pato, flamingo, gaivota, falcão (abutre), *kiwi* (tipo de ave), graçola, avestruz, periquito, pingüim, faisão, ema, gaivota negra, *skua* (ave marinha), pardal, cisne, águia, cambaxirra.

Classe 03: contém um conjunto de 5 instâncias (animais) - **répteis**
Conjunto de animais: *pitviper*, *seasnake*, *slowworm* (cobra de vidro, anguinha), tartaruga, tuatara (pequeno réptil).

Classe 04: contém um conjunto de 13 instâncias (animais) - **peixes**
Conjunto de animais: *bass*, carpa, peixe gato (lampreia), caboz (espécie de carpa), tubarão, *hadoque*, arenque, *pike* (lúcio), piranha, cavalo marinho (morsa), solha, *stingray*, atum.

Classe 05: contém um conjunto de 4 instâncias (animais) - **anfíbios**
Conjunto de animais: sapo, rã, salamandra (anfíbio), toad (outra espécie de sapo).

Classe 06: contém um conjunto de 8 instâncias (animais) - **insetos**
Conjunto de animais: pulga, mosquito (inseto), abelha, mosca de casa, joaninha, mariposa (traça), cupim, vespa.

Classe 07: contém um conjunto de 10 instâncias (animais) - **crustácios**
Conjunto de animais: molusco, caranguejo, camarão, lagosta, polvo, escorpião, *seawasp*, lesma, estrela do mar, verme.

Cada instância, ou seja, cada animal possui 18 atributos, isto é, 18 características físicas ou biológicas. Desses 18 atributos, 1 atributo é o nome do

animal, 15 atributos booleanos (1-verdadeiro, 0-falso) e 2 atributos numéricos. Na tabela a seguir estão definidos os atributos:

Número	**Descrição**	**Tipo**
01	nome do animal	Caracter
02	cabelo	Booleano
03	pena	Booleano
04	ovos	Booleano
05	leite	Booleano
06	aerotransportado	Booleano
07	aquático	Booleano
08	predador	Booleano
09	dentado	Booleano
10	coluna vertebral	Booleano
11	respira	Booleano
12	venenoso	Booleano
13	barbatana	Booleano
14	pernas	Lista de valores: (0, 2, 4, 5, 6, 8)
15	rabo	Booleano
16	doméstico	Booleano
17	felino	Booleano
18	tipo (classe a que pertence)	Intervalo de valores: (1,7)

Tabela 1: Atributos utilizados nos treinamentos - *Backpropagation*.

Analisando a tabela de atributos, pode-se notar que os atributos que estão entre os números 2 e 13 e entre os números 15 e 17 são booleanos: só podem assumir valores 0 (falso) ou 1 (verdadeiro), por meio dos quais se saberá se o animal possui (1) ou não possui (0) o atributo questionado.

Tomemos, como exemplo, o atributo 4, que informa se o animal bota ovo ou não. Para a pergunta da rede "Bota ovos?" diz-se 0 (não), caso o animal não bote ovo, e 1, em caso contrário. O atributo 3 informa se o animal tem penas ou não; durante o treinamento é feita a pergunta "Tem pena?", cuja resposta será 0, caso o animal não tenha pena, e 1, em caso contrário.

Já o atributo 1 somente é utilizado para definir o nome de cada animal correspondente aos demais atributos.

Como um atributo numérico, o de número 14 é utilizado para definir a quantidade de pernas que o animal possui, na lista de valores definida na tabela: 0 para quem não tem pernas, 2 para quem tem duas pernas, 4 para quem tem quatro pernas, 5 para quem tem cinco pernas, 6 para quem tem seis pernas ou 8 para quem tem oito pernas.

Serve o atributo 18 para classificar o animal entre as sete classes definidas anteriormente. Por conseguinte, a classificação dos animais é feita pelo atributo 18 e complementada pelos demais atributos.

4.3.2. Definição das camadas de entrada, intermediária e de saída

A *camada de entrada* será representada por 16 unidades de processamento (neurônios). Cada unidade representará um atributo entre o 2 e o 17, definido na tabela anterior e utilizado para informar à rede sobre a presença ou não do atributo no animal. Os atributos 1 e 18, portanto, não estão na camada de entrada, já que será por intermédio deles que a rede dará o resultado da classificação dos animais.

Na *camada intermediária*, pode-se ter uma variação quanto ao número de unidades de processamento. Inicialmente, estabelece-se um valor duas vezes menor que a quantidade de unidades de processamento da camada de entrada ou da quantidade de atributos. No exemplo, 16 atributos de entrada serão divididos por 2. O resultado apresentará 8 atributos, que novamente serão divididos por 2, obtendo-se 4 atributos que iniciarão o processamento na camada intermediária. A *camada de saída* possuirá apenas uma unidade de processamento, na qual será utilizada para a resposta da rede.

O treinamento tem seu início via interrogação da rede, isto é, fornece-se à rede informações sobre os animais utilizando-se os atributos 2 ao 17, representados pela camada de entrada.

4.3.3. Definição das Arquiteturas

Para os treinamentos, testes e validações da rede, foram utilizadas as seguintes *arquiteturas* (topologias):

Treinamento	Unidades de Processamento da Camada de Entrada	Unidades de Processamento da Camada Intermediária	Unidades de Processamento da Camada de Saída
1°	16	4	1
2°	16	5	1
3°	16	6	1

- **Taxa de Aprendizado:** 0.60 (60%); 0.70 (70%) e 0.90 (90%)
- **Taxa de Erro:** 0.05 (5%) e 0.10 (10%)

Tabela 2: Arquiteturas utilizadas nos treinamentos – *Backpropagation*.

Para as três arquiteturas definidas na tabela anterior, as *taxas de aprendizagem* utilizadas para estabelecer o valor no qual a resposta de saída deva

se aproximar do estímulo de entrada foram iniciadas em 0.60, 0.70 e 0.90. A *taxa de erro máximo aceitável* foi estabelecido em 0,05 para uma primeira bateria e de 0.10 para uma segunda bateria de simulações.

A *taxa de aprendizado* indica o percentual de acertos que a rede deve alcançar sobre o montante de atributos (16) apresentado à rede. A *taxa de erro* indica o percentual que a rede poderá errar referente ao montante de instâncias (101) apresentadas à rede.

Utilizando as três arquiteturas da tabela anterior, pode-se classificar qual delas obterá melhor resultado. Noutros termos, saberemos qual dessas arquiteturas terá mais acertos quanto à classificação dos animais.

4.3.4. Configurações das Arquiteturas

Para os treinamentos aqui propostos, serão utilizadas uma camada de entrada, uma camada intermediária e uma camada de saída. As variações nessas arquiteturas, portanto, ocorrerão no número de unidades de processamento da camada intermediária, no valor da taxa de aprendizagem e no valor da taxa de erro.

O número de unidades de processamento da camada intermediária utilizados foram os seguintes: 4 unidades para o primeiro treinamento, 5 para o segundo e 6 para o terceiro. A camada de entrada será fixada em 16 unidades de processamento e a de saída em 1 unidade de processamento, como ilustrado na Tabela 2.

- **Primeiro treinamento**
  - camada de entrada: 16 unidades de processamento;
  - camada intermediária: 4 unidades de processamento;
  - camada de saída: 1 unidade de processamento.

A Figura 23, de própria autoria, ilustra as configurações utilizadas neste treinamento. A rede contém 21 unidades de processamento e 68 conexões sinápticas.

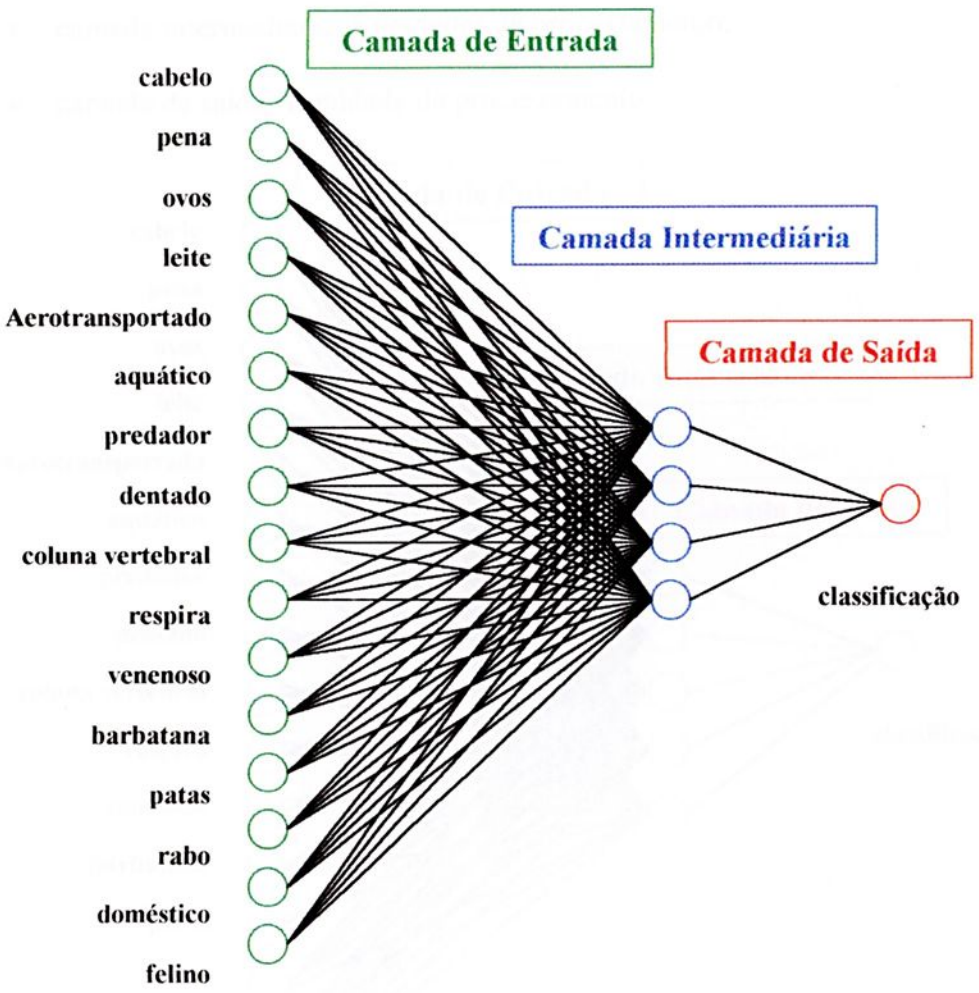

Figura 23: Configuração da rede para o primeiro treinamento- *Backpropagation*

Na tabela a seguir, encontram-se os resultados das simulações do primeiro treinamento:

Taxa de Aprendizagem	0.60	0.60	0.70	0.70	0.90	0.90
Taxa de Erro	0.05	0.10	0.05	0.10	0.05	0.10
% de acerto	**90**	**98-100**	**88**	**96**	**86**	**93**

Tabela 3: Resultados do primeiro treinamento.
Nota: o valor 98-100 indica que nesse fase do treinamento a porcentagem de acertos variou de 98 à 100%.

- Segundo treinamento
  - camada de entrada: 16 unidades de processamento;
  - camada intermediária: 5 unidades de processamento;
  - camada de saída: 1 unidade de processamento.

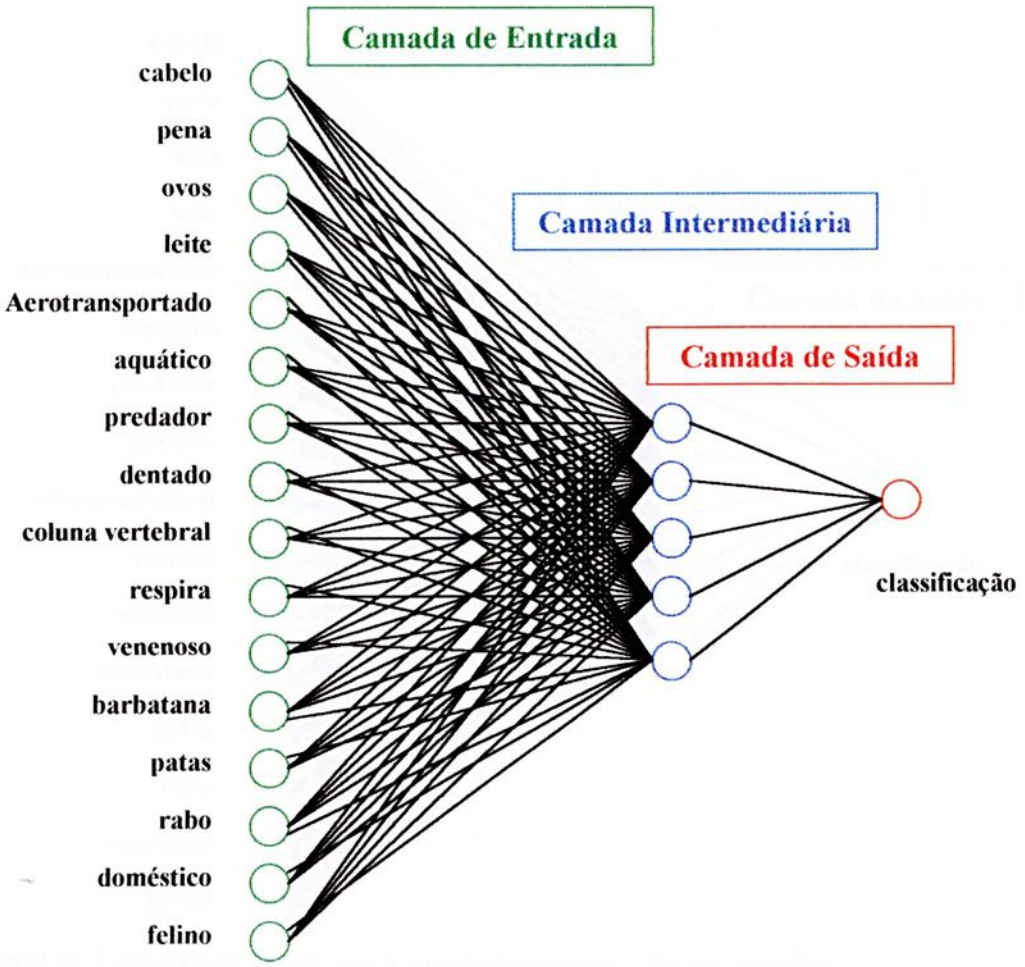

Figura 24: Configuração da rede para o segundo treinamento - *Backpropagation*.

A figura anterior, de própria autoria, ilustra as configurações utilizadas nesse treinamento. A rede contém 22 unidades de processamento, 85 conexões sinápticas.

Na tabela seguinte, se encontram os resultados das simulações do segundo treinamento:

Taxa de Aprendizagem	0.60	0.60	0.70	0.70	0.90	0.90
Taxa de Erro	0.05	0.10	0.05	0.10	0.05	0.10
% de acerto	**78**	**88**	**72**	**78**	**65**	**69**

Tabela 4: Resultados do segundo treinamento.

- **Terceiro treinamento**

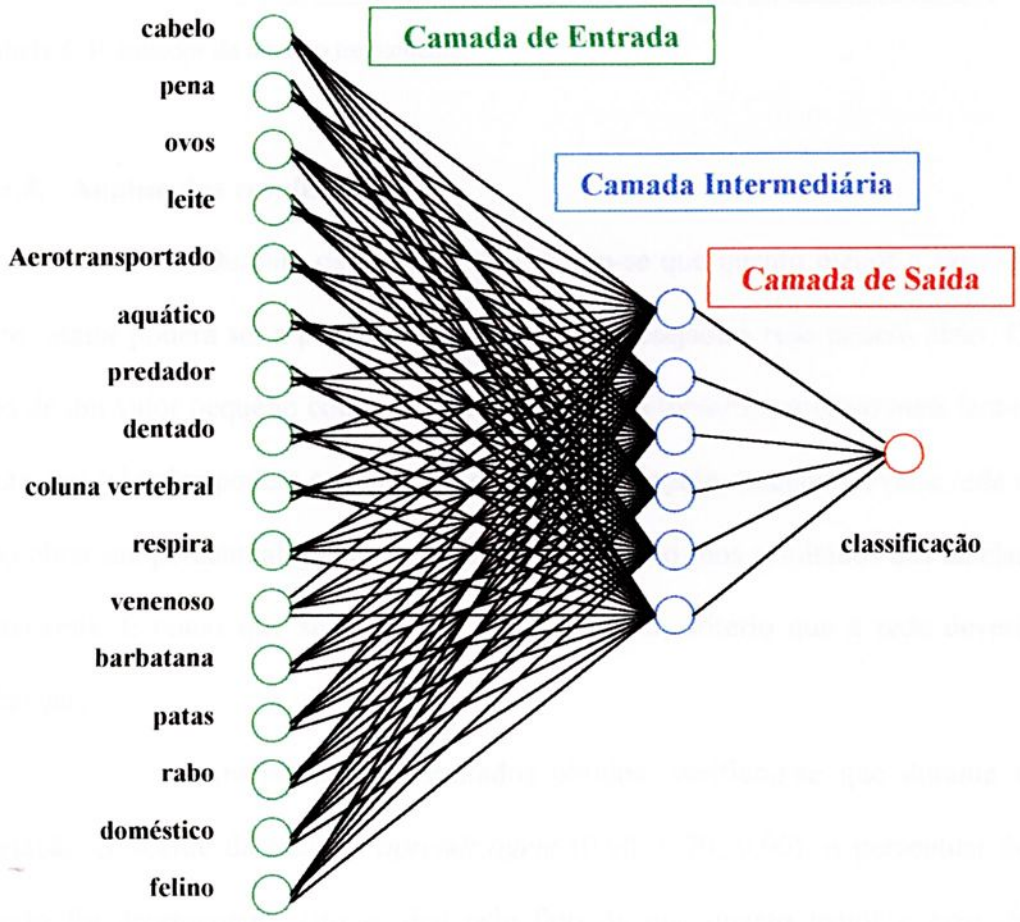

Figura 25: Configuração da rede para o terceiro treinamento – *Backpropagation*.

- camada de entrada: 16 unidades de processamento;
- camada intermediária: 6 unidades de processamento;
- camada de saída: 1 unidade de processamento.

A Figura 25, de própria autoria, ilustra as configurações utilizadas nesse treinamento. A rede contém 23 unidades de processamento, 102 conexões sinápticas.

Na tabela a seguir, se encontram os resultados das simulações do terceiro treinamento:

Taxa de Aprendizagem	0.60	0.60	0.70	0.70	0.90	0.90
Taxa de Erro	0.05	0.10	0.05	0.10	0.05	0.10
% de acerto	**72**	**81**	**67**	**68**	**60**	**65**

Tabela 5: Resultados do terceiro treinamento.

4.3.5. Análise dos resultados

Durante os treinamentos, notou-se que quanto menor a *taxa de erro*, maior poderá ser a proximidade da resposta desejada a rede poderá obter. O uso de um valor pequeno como *taxa de erro*, poderá tornar o processo mais lento, visto que tal valor poderá ser mais difícil de ser alcançado, podendo levar a rede a não obter um percentual grande de acertos, como visto nos resultados das tabelas anteriores. É como que se a *taxa de erro* fosse um critério que a rede deverá alcançar.

Analisando os resultados obtidos, verificou-se que durante a variação crescente da *taxa de aprendizagem* (0.60; 0.70; 0.90), o percentual de acerto foi decrescendo. Isto se deu pelo fato de que quanto maior a *taxa de*

aprendizagem, menos adaptações de parâmetros serão realizados no algoritmo, levando a rede a ficar mais longe da saída desejada. Portanto, quando menor a *taxa de aprendizagem* mais perto da saída desejada a rede tentará chegar. É como se a taxa de aprendizagem fosse um critério estabelecido para a rede alcançar a saída mais próxima da entrada.

Divididos pela taxa de erro, os resultados foram os seguintes:

Taxa de erro 0.05:

Arquitetura	Taxa Aprendizagem 0.60	Taxa Aprendizagem 0.70	Taxa Aprendizagem 0.90
16-4-1	90	88	86
16-5-1	78	72	65
16-6-1	72	67	60

Tabela 6: Resultados dos treinamentos - taxa de erro 0.05– *Backpropagation.*

Taxa de erro 0.10:

Arquitetura	Taxa Aprendizagem 0.60	Taxa Aprendizagem 0.70	Taxa Aprendizagem 0.90
16-4-1	98-100	96	93
16-5-1	88	78	69
16-6-1	81	68	65

Tabela 7: Resultados dos treinamentos - taxa de erro 0.10– *Backpropagation.*

Considerando as tabelas dos resultados, percebe-se também que a arquitetura que melhor se adaptou em relação ao problema proposto foi a que utilizou 4 unidades de processamento na camada intermediária, com taxa de aprendizagem 0.60 e taxa de erro 0.10. Isso não implica que, em termos de *quantidade exata*, esta arquitetura acertou mais do que a arquitetura que utilizou a taxa de erro 0.05. Isto se deve a que, quanto menor a taxa de erro, mais exata deverá ser a resposta. Em caso contrário, a rede não acertará. Portanto, pode-se notar que, quanto maior a taxa de erro, maior poderá ser o *percentual* de acerto,

visto que com uma taxa de erro alta, a possibilidade de aproximação exata não é assim tão necessária, já que a taxa de erro proporciona o acerto. Por outro lado, menor poderá ser a *quantidade exata (real)* de acertos quanto menor a taxa de erro. Ou seja, a taxa de erro alta não força uma aproximação exata do resultado. Disso se infere que quanto menor a taxa de erro, menor poderá ser o *percentual* de acerto, visto que com uma taxa de erro baixa, a possibilidade de aproximação exata é necessária já que a taxa de erro exige o acerto exato. Por outro lado, maior poderá ser a *quantidade exata* de acertos, ou seja, a taxa de erro baixa força uma aproximação exata do resultado.

É claro que, para a conclusão anterior, não se pode deixar de lado a taxa de aprendizagem, visto que é por intermédio dela que a rede tenta atender ao valor da taxa de erro exigido. Nem sempre, porém, a taxa de aprendizagem consegue tal ação.

Um outro ponto importante a ser relatado aqui é quanto ao número de unidades de processamento na camada intermediária. Verificou-se que o número de unidades de processamento ideal para tal camada deverá ser um valor duas vezes menor do que o número de unidades de processamento da camada de entrada. Portanto, a camada intermediária que melhor se adaptou, foi a que continha quatro unidades de processamento. Ao aumentar o número de unidades de processamento da camada intermediária, mantendo as mesmas taxas de aprendizagem e de erro, verificou-se que o percentual de acertos foi diminuindo. Isso é decorrente do fato de que o número de unidades de processamento da camada intermediária deve ser calculado de acordo com os critérios exigidos pelo

modelo em uso e pela quantidade de unidades de processamento da camada de entrada. Não se trata de diminuir ou aumentar a quantidade de unidades a gosto.

O fato de simplesmente diminuir ou aumentar o número de unidades de processamento da camada intermediária não quer dizer que será melhor ou pior. Portanto, definido o número de unidades de processamento da camada intermediária, deve-se treinar a rede variando somente as taxas de aprendizagem e de erro. Assim, busca-se uma melhor performance da rede.

Deve-se salientar, entretanto, que foi utilizada apenas uma camada intermediária, mas para testes futuros, deverão ser utilizadas mais camadas intermediárias, a fim de confrontar os resultados aqui obtidos.

4.4. Treinamento de uma Rede Neural Artificial *SOFM* de Kohonen para reconhecimento e classificação de padrões informacionais

Apresenta-se aqui uma rede cujo objetivo é detectar (até mesmo criar) e classificar padrões de entradas informacionais. De acordo com as características desse padrão de entrada, a rede deverá gerar como saída outro padrão que se aproxime ao máximo do padrão de entrada apresentado à rede.

O software utilizado é de autoria de Fábio Luiz Viana, orientado pelo professor Dr. Norian Marranghello, da Universidade Estadual Paulista, Câmpus de São José do Rio Preto, do Instituto de Biociências, Letras e Ciências Exatas, Departamento de Ciências da Computação e Estatística (1994).

4.4.1. Configuração da rede para as simulações

A *camada de entrada* e a *camada de saída* serão representadas, cada uma, por cinco unidades de processamento (neurônios), compondo um total de vinte e cinco unidades de processamento.

Como *raio de vizinhança*, o software nos fornece três opções:

- *raio definido como sendo o próprio tamanho da rede*: esse raio abrange todas as unidades de processamento da rede, levando todas as unidades a participarem da competição para saber qual unidade sairá vencedora;
- *raio fixado em 75% do tamanho da rede*: esse raio abrange uma faixa de 75% das unidades de processamento da rede, deixando 25% fora da competição;
- *raio definido pelo usuário da rede*: é raio que poderá abranger um tamanho escolhido pelo usuário. O que se alerta aqui é que quanto menor for o raio, menos unidades de processamento competirão tendo implicações, assim, nos resultados finais das simulações.

Nota-se, portanto, que quanto maior o raio de vizinhança, melhor poderá ser o resultado obtido, visto que, uma maior quantidade de unidades de processamento estarão participando da competição para se tornarem a unidade vencedora. Nesse treinamento utilizou-se, durante as simulações, um raio de vizinhança fixado em 75% em relação ao tamanho da rede.

Como *taxa de aprendizagem* o software nos dá a opção de escolha entre os valores 0.1 a 0.9. É bom salientar que quanto menor a taxa de aprendizagem, mais perto da resposta de saída a resposta obtida poderá chegar. Por outro lado, o treinamento levará muito mais tempo para ser concluído. Assim

optou-se aqui por uma taxa de aprendizagem não tão baixa, mas também não tão alta. O valor da taxa foi fixado em 0.6.

Como *padrão de entrada*, o software nos fornece cinco opções de escolha: *distribuição uniforme*, *quadrado*, *círculo*, *anel* ou *triângulo*. Nesse treinamento optou-se pelo *quadrado* como padrão de entrada.

Quanto à *quantidade de iterações* entre as unidades de processamento, o software nos fornece a opção de escolha entre os valores 1 e 3000 iterações. A quantidade de iterações também auxilia na obtenção de um melhor resultado. Com uma maior quantidade de iterações, as unidades de processamento terão mais opções de escolha entre elas, isto é, trocarão mais informações, se comunicarão muito mais antes de resultar-se a unidade vencedora. Portanto, quanto maior for a quantidade de iterações, mais próxima do padrão de entrada fornecido a rede chegará.

4.4.2. O Treinamento

Para esse treinamento, a rede teve a seguinte configuração: vinte e cinco unidades de processamento, raio de vizinhança fixado em 75%, taxa de aprendizagem inicial 0.6, sendo o padrão de entrada um quadrado. Quanto à quantidade de iterações, utilizou-se três valores: 1000 iterações para a primeira simulação, 2000 para a segunda e 3000 para a terceira.

Configuração inicial da rede:

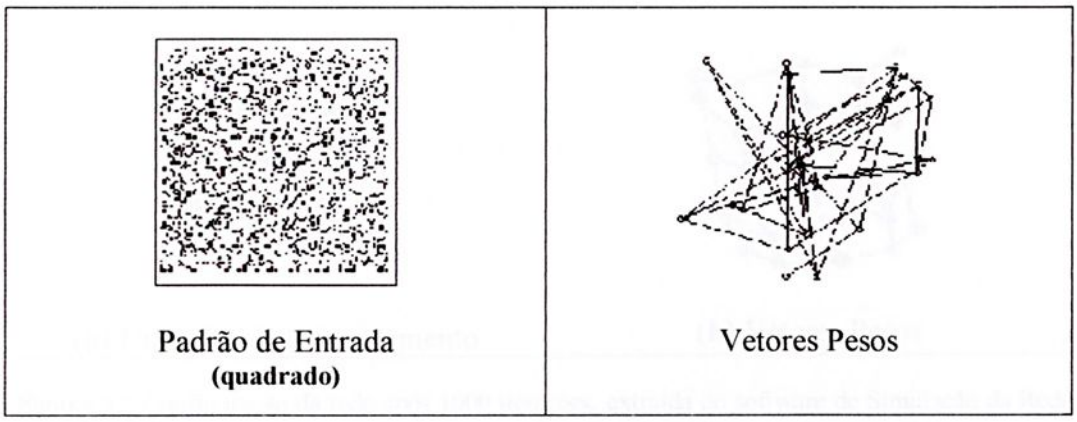

Figura 26: Configuração inicial da rede para o treinamento, extraída do software de Simulação da Rede Neural de Kohonen.

Na parte (a) das Figuras 27, 28 e 29, tem-se a distribuição das unidades de processamento. As unidades de cor azul são as unidades derrotadas (inativas) durante as simulações; as unidades de cor verde são as que ficaram dentro do raio de vizinhança; e as unidades de cor vermelha foram as unidades vencedoras (ativas).

Na parte (b) das Figuras 27, 28 e 29, encontram-se os resultados obtidos após cada simulação, de acordo com a quantidade de iterações (1000; 2000; 3000).

Após 1000 iterações entre as unidades de processamento, envolvidas na primeira simulação, a configuração compôs-se da seguinte forma:

- unidade vencedora: 3
- taxa de aprendizagem: 0.0006
- raio de vizinhança: 1

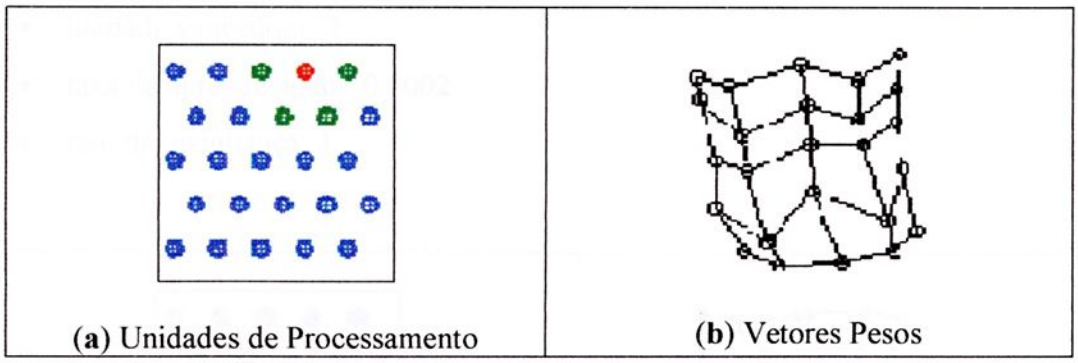

Figura 27: Configuração da rede após 1000 iterações, extraída do software de Simulação da Rede Neural de Kohonen.

Após 2000 iterações entre as unidades de processamento, envolvidas na segunda simulação do treinamento, a configuração compôs-se da seguinte forma:

- unidade vencedora: 19
- taxa de aprendizagem: 0.0003
- raio de vizinhança: 1

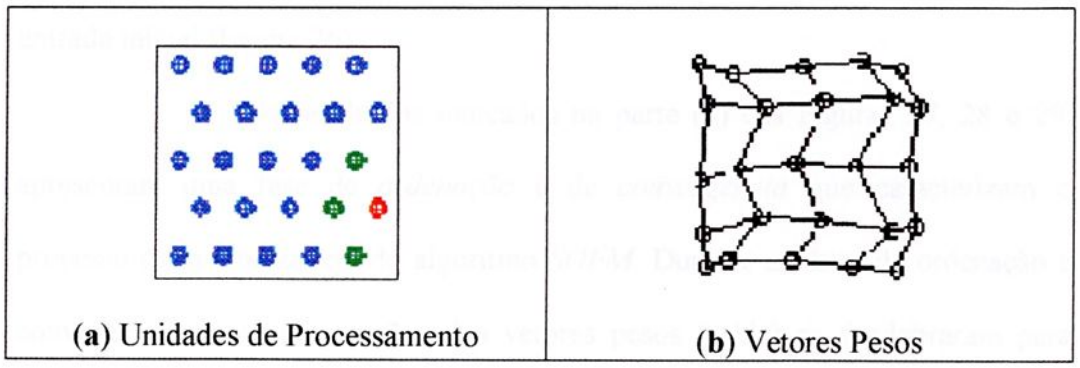

Figura 28: Configuração da rede após 2000 iterações, extraída do software de Simulação da Rede Neural de Kohonen.

Após 3000 iterações entre as unidades de processamento, envolvidas na segunda simulação do treinamento, a configuração compôs-se da seguinte forma:

- unidade vencedora: 2
- taxa de aprendizagem: 0.0002
- raio de vizinhança: 1

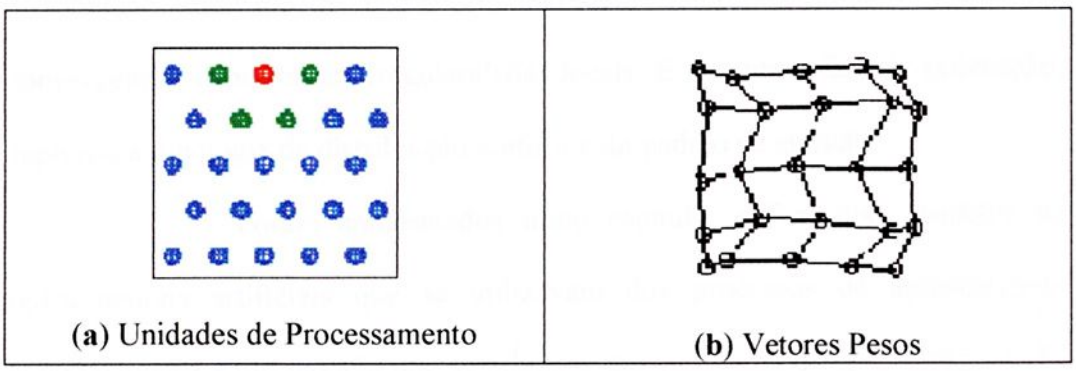

Figura 29: Configuração da rede após 3000 iterações, extraída do software de Simulação da Rede Neural de Kohonen.

Nota-se que a cada simulação, com o aumento da quantidade de iterações entre as unidades de processamento, o resultado obtido, ilustrado na parte (b) das Figuras 27, 28 e 29, foi se aproximando cada vez mais do padrão de entrada inicial (Figura 26).

Os resultados indicados na parte (b) das Figuras 27, 28 e 29, apresentam uma fase de *ordenação* e de *convergência* que caracterizam o processo de aprendizagem do algoritmo *SOFM*. Durante as fases de ordenação e convergência, as configurações dos vetores pesos (saída) se desdobraram para formarem malhas e se estenderem para preencher o espaço do padrão inicial.

Durante a fase de ordenação, verificou-se que as unidades de processamento foram mapeadas na ordem correta do padrão de entrada inicial. Na a fase de convergência, a distribuição estatística das unidades de processamento

foi detectada relativamente próxima à distribuição estatística do padrão de entrada inicial, exceto por alguns efeitos em torno do resultado final.

Verificou-se, pois, que o resultado final dos vetores pesos com a distribuição uniforme do padrão de entrada (Figura 26), durante a fase de convergência, capturou as irregularidades locais. E durante a fase de ordenação, capturou a topologia da distribuição uniforme do padrão de entrada.

Foram apresentados neste capítulo, enfim, dois modelos de redes neurais artificiais que se utilizavam dos processos de aprendizagem supervisionada e um modelo que se utilizava dos processos de aprendizagem não supervisionada. Como modelos de redes neurais artificiais com aprendizagem supervisionada (organização), vimos os modelos *perceptron de camada única* e o *perceptron de múltiplas camadas*, este último conhecido por algoritmo de *backpropagation*. Como modelo de redes neurais artificiais com aprendizagem não supervisionada, estudamos o modelo *SOFM* de Kohonen. Dois treinamentos para reconhecimento e classificação de padrões informacionais forma considerados. Para as simulações do primeiro treinamento, utilizou-se o modelo *perceptron de múltiplas camadas* com aprendizagem supervisionada; e nas simulações do segundo treinamento, utilizou-se o modelo *SOFM* de Kohonen, com aprendizagem não supervisionada.

Os resultados aqui obtidos, serão essenciais para o desenvolvimento do último capítulo e das considerações finais.

A seguir procederemos a uma análise dos modelos de redes neurais artificiais aqui apresentados, em relação aos processos de auto-organização.

5

Análise do Modelo *Perceptron de Múltiplas Camadas* (*Backpropagation*) e do Modelo *SOFM* de Kohonen em relação a Processos de Auto-organização

Após embasamento adquirido na elaboração dos capítulos 1 a 4 nos quais foram analisados conceitos e fundamentos da *Auto-organização*, de *Redes Neurais Artificiais*, de *Métodos de Aprendizagem* e de *Modelos de Redes Neurais Artificiais com Aprendizagem Supervisionada* e *Não Supervisionada* considera-se possível discutir o assunto central desse trabalho: *verificar em que medida processos de auto-organização ocorrem nos modelos de redes neurais artificiais, por meio dos métodos de aprendizagem supervisionada e não supervisionada.*

Segundo Atlan (1992), o que se busca é averiguar o fenômeno da auto-organização junto aos modelos de redes neurais artificiais, ou seja, verificar se os sistemas organizados evoluem ou se auto-organizam, e de que maneira isso ocorre.

Pode-se conceber a evolução de sistemas organizados, ou o fenômeno de auto-organização de acordo com Atlan (1992), como,

> um processo de aumento de complexidade, simultaneamente estrutural e funcional, resultante de uma sucessão de desorganizações resgatadas, acompanhadas, em todas as ocasiões, pelo restabelecimento num nível de variedade maior e de redundância mais baixa.

Assim, na concepção de Atlan (1992, p. 45), o estado de organização de um sistema é definido não apenas por sua quantidade de informação, mas também por sua organização funcional descrita pela taxa de variedade da quantidade de informação presente no sistema.

Analisando-se esta concepção, nota-se que ela se encaixa perfeitamente nos modelos de redes neurais artificiais, visto que os modelos de redes neurais artificiais se utilizam, muitas vezes, durante suas simulações (treinamentos), de uma grande variedade de informação. Quando se tem uma grande variedade de opções na escolha da informação a ser recebida, certamente será possível obter-se um número muito maior de informações que mais se aproximam daquilo que se deseja. Por outro lado, essa variedade de informação poderá causar uma diminuição de redundância, ocasionando um leque mais variado de informações, sem duplicação. Além disso, os modelos de redes neurais artificiais também possuem uma *grande complexidade estrutural e funcional*, visto que se pode trabalhar com um número variado de unidades de processamento. Sendo assim, há um número muito maior de conexões entre essas unidades, o que torna o sistema neural artificial bem mais complexo.

O propósito central dessa análise é *verificar se os sistemas neurais artificiais evoluem ou se auto-organizam durante seu funcionamento* (treinamento em execução). Para tanto, dividiu-se esse capítulo em duas seções. A primeira, tratará do método de aprendizagem supervisionada; a segunda, do método de aprendizagem não supervisionada.

5.1. O modelo *Perceptron de Múltiplas Camadas – Backpropagation* (aprendizagem supervisionada)

Segundo Gonzales (1996), os modelos de redes neurais artificiais com aprendizagem supervisionada não são considerados modelos que se auto-organizam, visto que dependem, direta e indiretamente, da presença de um supervisor para controlar todas as etapas do processo de estruturação e desenvolvimento da rede. Como foi visto na subseção 3.2.1, é o supervisor quem fornece à rede os estímulos de entrada juntamente com um estímulo de saída desejado. Após o treinamento, espera-se que a rede encontre o estímulo de saída que melhor se associou ao estímulo de entrada.

Na prática, verificou-se durante as simulações realizadas para reconhecimento e classificação de características de padrões animais, que todos os estímulos de entrada e saída foram fornecidos *a priori*. Essas simulações se deram da seguinte forma: foi fornecida à rede uma planilha (vetor de informações) contendo, em cada linha, características corretas de cada animal, definidas como estímulos de saída (resposta desejada); também foi dada à rede uma outra planilha, contendo características para cada animal, definidas como estímulos de entrada. O papel dessa segunda planilha era interrogar a rede de acordo com as respostas da primeira.

Em seguida, a rede foi abastecida com a taxa de aprendizagem e a taxa de erro aceitável, e depois foi colocada em treinamento. Durante esse processo, a rede comparou os estímulos de entrada aos estímulos de saída de cada animal, fazendo uma contagem dos acertos obtidos de acordo com a taxa de

aprendizagem. Se a quantidade de acertos atendesse ao que foi delimitado como taxa de erro aceitável, a rede acertaria o animal em questão. Caso contrário, um erro era propagado de volta à rede para recálculo, a fim de tentar deixá-lo abaixo da taxa de erro aceitável. É bom esclarecer que os estímulos de entrada e de saída ficaram a critério do supervisor, que pode tanto aproximar ou distanciar as entradas das saídas desejadas. Após algum tempo, a rede apresenta a configuração final do que ela conseguiu acertar em relação ao conjunto de animais em questão.

Por conseguinte, em treinamentos que se utilizam do método de aprendizagem supervisionada, a presença do supervisor, que fornece todos os estímulos de entrada e de saída *a priori*, é indispensável, caso contrário a rede nada faz.

Em se tratando do método de aprendizagem supervisionada, não são encontrados, em síntese, processos de auto-organização. Esse método não apresenta qualquer característica ou propriedade de auto-organização, sendo dependente única e exclusivamente de um supervisor. Essas redes são utilizadas simplesmente para reconhecimento e classificação de padrões informacionais, nada mais que isso, tudo sendo determinado pelo supervisor e pelo algoritmo.

Após um vasto embasamento sobre o método de aprendizagem supervisionado, pôde-se, enfim, levantar alguns pontos que o definem como um método no qual nada se encontra em relação aos processos de auto-organização. O primeiro ponto a considerar diz respeito à presença necessária de um supervisor, ou seja, é por intermédio do supervisor que a rede recebe os estímulos de entrada, os quais podem ser provenientes do meio ambiente externo. O possuidor do

conhecimento sobre o ambiente em questão é o supervisor. Entretanto, esse ambiente é desconhecido pela rede neural em questão. Um segundo ponto a relatar refere-se à resposta desejada. Em virtude de seu conhecimento prévio, o supervisor é quem fornece à rede uma resposta desejada, em que representa a ação ótima a ser alcançada pela rede para aquele treinamento em questão. Um último ponto aqui mencionado, talvez um dos mais importantes, é o fato de como se dá o treinamento da rede. Após a bateria de treinamentos realizados na seção 4.3, pôde-se inferir que o que realmente ocorre, durante os treinamentos, são comparações entre os estímulos de entrada e o estímulo de saída desejado. Nesses estímulos tem-se um percentual de erros definido pela taxa de erro e uma função de cálculo dos pesos, caso o percentual de erros não se encontre abaixo da taxa de erro aceitável. A função de cálculo dos pesos é que realmente tenta, por meio da correção dos erros (taxa de aprendizagem), fazer adaptações dos procedimentos e funções (parâmetros) do algoritmo para que esses erros fiquem abaixo da taxa de erro aceitável, fixada no inicio do treinamento. Portanto, aqui não ocorre nenhum processo de auto-organização, tudo está definido e fixado no algoritmo por meio dos procedimentos e funções pré-elaborados, que podem sofrer pequenas adaptações em seus parâmetros.

Tais constatações, já eram pressupostas, considerando-se que, por definição, esses modelos possuem a presença de um supervisor.

5.2. O modelo *SOFM* de Kohonen (aprendizagem não supervisionada)

O modelo *SOFM* de Kohonen deve, por si próprio, detectar e classificar os estímulos de entrada e, se necessário, criar tais estímulos por meio do reconhecimento de regularidade na informação disponível no meio ambiente externo.

Como já se afirmou, nesse método não existe a presença de um supervisor; quem faz as vezes do supervisor é o meio ambiente externo, juntamente com o algoritmo. O que realmente acontece é uma contínua interação da rede com o meio ambiente externo, levando aquela a receber informações para sua alimentação ou realimentação.

Na prática, os estímulos de entrada são recebidos pela rede, vindos do meio ambiente externo, por intermédio de mecanismos eletrônicos, como por exemplo os sensores magnéticos (sinais sensoriais), ligados às unidades de processamento da camada de entrada.

Pela contínua interação com o meio ambiente externo, os estímulos de entrada são recebidos inúmeras vezes, graças à presença da redundância. Desta forma, as muitas entradas podem ser reconhecidas e classificadas. A partir daí é que a rede começará a gerar estímulos de saída.

Portanto, os estímulos de saída também não são fornecidos à rede. Ela própria gera os estímulos de saída por meio da redundância dos estímulos de entrada por ela recebidos.

Falar de meio ambiente externo não significa dizer que nele se configure todo o universo. Somente interessa aqui o ambiente relacionado ao ciclo

específico que se deseja pesquisar. No exemplo de treinamento, dado na seção 4.3, pretendia-se reconhecer e classificar características de animais. O ambiente externo se posicionou por um conjunto de animais específicos.

Pode-se aplicar o modelo *SOFM* de Kohonen em vários ciclos (situações) diferentes, ou seja, em vários ambientes diferentes. O que deverá servir de ambiente externo, na pesquisa em discussão, será o próprio ambiente em questão. A interferência do ruído de uma maneira informacional fará a diferença, podendo dificultar ou colaborar para o reconhecimento e classificação mais adequado ao estímulo de entrada recebido pela rede.

Para melhor compreensão dos mecanismos que estão sendo analisados, a auto-organização é vista sob duas perspectivas que envolvem os modelos utilizadores do método de aprendizagem não supervisionada.

A primeira perspectiva é a *auto-organização primária*. Na concepção de Ashby (1962), podem-se citar exemplos de esquemas, como eco-sistema, ser humano. São sistemas composto de subsistemas, que por sua vez também possuem subsistemas, e assim sucessivamente. Nesse exemplo, confirma-se a aplicação do conceito de que partes separadas e incondicionadas possam vir a formar partes unidas, condicionadas e, enfim, organizadas, compondo o todo. Ashby também define esse primeiro tipo de auto-organização como simples "auto conexão" entre elementos.

O exemplo dado na subseção 1.2.4., sobre formação de um grupo de pesquisa, ilustra um processo prático de auto-organização primária, como definido no parágrafo anterior.

cm 1 2 3 4 5 6 7 unesp 1 12 13 14 15 16 17 18

Ainda em se, aplicando o conceito de partes e todo de Ashby aos modelos *SOFM* de Kohonen, as partes podem ser consideradas como as unidades de processamento que inicialmente se encontram conectadas, de acordo com a arquitetura utilizada (ver seção 2.5), mas sem qualquer relação de dependência entre elas, isto é, unidades de processamento independentes. A partir do momento em que se inicia o treinamento, a rede começa a interagir com o meio ambiente externo, desencadeando uma propagação entre as unidades de processamento, por meio de estados de ativação excitatórios ou inibitórios. A rede eventualmente convergirá, criando uma estrutura ou forma organizada, que se pode considerar como o todo. Diz-se, então, que a rede se auto-organizou, visto que uma estrutura (todo) foi criada, primariamente, a partir de elementos inicialmente independentes, sem a intervenção de um supervisor.

A segunda perspectiva é a *auto-organização secundária*, na qual se pressupõe a existência de um sistema primariamente organizado, adicionando-se nela mecanismos de evolução. A auto-organização secundária inicia-se de um sistema inicialmente criado pela auto-organização primária, cujo objetivo é tentar levar esse sistema a evoluir, tornando-o um sistema auto-organizado secundariamente.

Segundo Gonzales (1996), baseando-se em Ashby (1962) e Debrun (1996), esses mecanismos de evolução incluem *feedback* (retroalimentação) e métodos de aprendizagem por meio de *correção de erros*. Acrescentam-se também *critérios de relevância* utilizados para distinguir as

situações boas das más e as situações adequadas das inadequadas para o desenvolvimento do sistema.

Ashby (1962) acredita que nenhuma máquina pode, em princípio, se auto-organizar secundariamente, visto que falta a ela a capacidade básica de desenvolver critérios de relevância, critérios esses, presentes nos sistemas vivos. Segundo Gonzales (1996, pp. 283-287), critérios de relevância são aqueles que ajustam os estados de equilíbrio dos sistemas vivos, como por exemplo: seleção de padrões informacionais e distinção entre ruído e outras formas de variedade de informação, preservando-os vivos.

A aquisição de critérios de relevância aparentemente não se reduz à simples redundância de informações provenientes do meio ambiente externo, mas também de uma habilidade inteligente de poder selecionar e distinguir o que realmente é necessário naquela situação em questão.

Com essa colocação de Ashby, pode-se, então, descartar a possibilidade de processos de auto-organização secundária nos sentidos forte e completo, visto que faltam às redes neurais artificiais, com aprendizagem não supervisionada, a capacidade de desenvolver critérios de relevância. Por outro lado, sabe-se que essas redes possuem a propriedade "*aprendizagem*". Tal mecanismo, associado à habilidade de desenvolvimento e aprimoramento dos seus estados representacionais, mesmo sem a capacidade de desenvolver critérios de relevância, pode gerar expectativa de auto-organização, pelo menos num sentido mais fraco, isto é, sem a presença da capacidade de desenvolver critérios de relevância.

cm 1 2 3 4 5 6 7 unesp 1 12 13 14 15 16 17 18

Para se confirmar a existência de processos de auto-organização secundária nas redes neurais artificiais, deve-se investigar se os métodos de aprendizagem presentes reúnem condições suficientes para uma possível evolução da rede.

Segundo Gonzales (1996, p. 285), duas situações criam a possibilidade de evolução ou aprimoramento em uma rede neural artificial: as que requerem *correção de erro* por meio de *feedback*; e as que empregam o *princípio da ordem a partir do ruído*.

A primeira possibilidade de evolução, que requer *correção de erros* por meio de *feedback*, aplica-se ao modelo *SOFM* de Kohonen que se utiliza do mecanismo de *correção de erros*. Quando o percentual de erros não fica abaixo da taxa de erro aceitável, o mesmo é propagado da camada de saída de volta à camada de entrada (de volta à rede). Neste momento, os pesos são recalculados, adaptando-se (ajustando) aos parâmetros da rede, na tentativa de levar tal percentual a tornar-se um valor abaixo da taxa de erro aceitável.

Nesse momento, após uma exposição redundante de *feedback* e com a *correção de erros*, a rede eventualmente terá conseguido formar padrões de informações especificadas pelo meio ambiente externo. Com isso, a rede, que inicialmente não fornecia respostas apropriadas aos estímulos de entrada provenientes do meio ambiente externo, passou a identificá-los, mas podendo possuir pequenas distorções (efeitos), como visto no treinamento realizado na Seção 4.4. Por outro lado, pode-se dizer que a rede aprendeu a corrigir e aprimorar

padrões informacionais. Entende-se que, nesse sentido, as redes neurais compartilham uma das propriedades da auto-organização secundária.

A segunda possibilidade de evolução emprega o *princípio da ordem a partir do ruído*. Nesse caso, sistemas abertos, submetidos a vários tipos de perturbações, podem vir a desenvolver mecanismos de adaptação, introduzindo variedade às suas formas fixas e bem estabelecidas de organização.

Logo, os sistemas auto-organizadores não se alimentam apenas da informação ordenada, podendo utilizar-se do ruído como outra fonte de alimentação. O que influencia muito nesta segunda possibilidade de evolução é a falta de desenvolvimento de critérios de relevância, visto que é por meio desses critérios que a rede poderia identificar a parte boa do ruído (que poderia contribuir para a evolução do sistema) e a parte ruim (a ser descartada pela rede, para não atrapalhar na evolução).

Por isso, fica claro que a presença do ruído em certos sistemas poderá ser caracterizada como algo bom, construtivo e não somente algo que interfira nos sistemas abertos de uma maneira ruim e destrutiva. Pode-se concluir que o ruído não somente é visto como um destruidor de sinais, mas também, como uma fonte de alimentação. Em se tratando de sistemas abertos ao meio ambiente, receptores de informações boas ou ruims, pode-se dizer que nesse momento está ocorrendo, por intermédio da auto-organização secundária, uma evolução no sistema.

Ao fim deste capítulo, salientamos que trouxemos uma análise dos modelos com *aprendizagem supervisionada* e do modelo com *aprendizagem*

não supervisionada, vistos no capítulo 4. Foi neste capítulo que se examinou em que medida processos auto-organizados poderiam ocorrer nos modelos de redes neurais artificiais com aprendizagem supervisionada e não supervisionada.

A seguir, serão feitas algumas considerações finais em relação ao objetivo inicialmente proposto.

CONSIDERAÇÕES FINAIS

Após o embasamento adquirido durante essa pesquisa, procederemos ao relato dos resultados obtidos em relação ao nosso objetivo, bem como à elaboração de algumas projeções para o futuro.

No início dessa pesquisa, pretendia-se averiguar se de fato os métodos de aprendizagem supervisionada e não supervisionada apresentavam propriedades em relação a processos da auto-organização, assim como investigar em que medida os modelos de redes neurais artificiais seriam capazes de modelar e de simular os fundamentos relativos à auto-organização.

Partiu-se, portanto, da idéia de Debrun (1996) de que a auto-organização tem sido caracterizada de maneira diferente e que, por trás dessas definições, existe uma idéia comum de que novas estruturas possam emergir, mais ou menos, espontaneamente em certos domínios.

Com base nesta concepção da auto-organização, deu-se início à presente investigação, tendo como foco de análise os modelos de redes neurais artificiais, com aprendizagem supervisionada e não supervisionada.

Supunha-se que esses dois tipos de aprendizagem eram utilizados por vários modelos de redes neurais artificiais, mas pretendia-se esclarecer até que ponto esses processos estavam incorporados, ou seja, de fato implantados nos modelos de redes neurais artificiais.

Tal pesquisa dividiu-se em duas partes: na primeira parte foram investigados modelos utilizadores do método de aprendizagem supervisionada. Trabalhou-se nesta parte com os modelos *Perceptron de Camada Única* e do modelo *Perceptron de Múltiplas Camadas* (*Backpropagation*). Na segunda parte foram estudados modelos utilizadores do método de aprendizagem não supervisionada. Trabalhou-se, nesta parte, com o modelo *SOFM* de Kohonen.

Para tal investigação, tornou-se primordial o estudo de alguns fundamentos teóricos e conceitos básicos que estavam diretamente ligados ao propósito global do trabalho. Portanto, foram introduzidos fundamentos e noções básicas sobre *Auto-organização*, *Sistemas Auto-organizados*, *Redes Neurais Artificiais*, *Métodos de Aprendizagem*, *Modelos de Redes Perceptron de Camada Única* e de *Múltiplas Camadas (Backpropagation)* e do *Modelo SOFM* de Kohonen (*Mapas de Características Auto-organizantes)*.

Portanto, pode-se considerar, com base na Seção 1.7 do Capítulo 1, que os modelos de redes neurais artificiais que se utilizam do método de aprendizagem supervisionado, aqui investigados, se encaixam nos processos que ficam *excluídos da auto-organização*, visto que são processos que dependem de um algoritmo no qual tudo já foi definido por um programador, e de um supervisor que comandará o processo.

Já os modelos de redes neurais artificiais que se utilizam do método de aprendizagem não supervisionada, aqui investigados, se encaixam no *grau zero* da auto-organização, visto que sofrem ajustes e aprimoramentos nos parâmetros de seu algoritmo, por meio a uma corrente interação com o meio

ambiente externo, além de se encaixarem no *grau 1*, em razão da existência de evolução de uma situação, o que se dá exclusivamente por meio ao autofuncionamento cibernético.

Os outros dois graus posteriores ao *grau 1* ficam fora do alcance das redes neurais artificiais, já que, necessariamente, se trata de sistemas humanos e não artificiais.

A propósito, vale salientar a nossa discordância em relação à definição de Atlan (1992, pp. 38-40) de que a ocorrência de auto-organização via ruído pode ser possível por meio de um *programa preciso* injetado no sistema. Essa situação se encaixa, sim, nos *processos que ficam excluídos da auto-organização*, visto que, por tratar-se de um programa preciso, não poderiam nele, ocorrer ajustes e aprimoramentos.

Vale salientar aqui também, a nossa discordância em relação a Kovács (1996) de que a rede neural artificial gera seu próprio algoritmo. Na realidade, a rede neural artificial apenas adapta o algoritmo via parâmetros por meio a estímulos recebidos pela rede neural artificial, e não o cria (gera) como cita Kovács (1996).

Finalizando, pode-se dizer que as redes neurais artificiais aqui investigadas, constituídas por um fluxo informacional, praticamente não possuem auto-organização. Assim, muitos ajustes, principalmente em relação a *critérios de relevância*, ainda serão necessários para que se possa evoluir nesta área.

cm 1 2 3 4 5 6 7 unesp 1 12 13 14 15 16 17 18

PROJEÇÕES FUTURAS

Atualmente, já se encontram disponíveis para comercialização diversos circuitos integrados que implementam redes neurais artificiais. Dentre os circuitos, podem-se citar os digitais, analógicos e híbridos, os quais implementam diversas arquiteturas de redes, com maior ou menor grau de interconexões, facilidade para interfaceamento, blocos de treinamento e outras características.

Segundo Yoneyama e Nascimento (2000, p. 167), as redes neurais artificiais podem ser utilizadas hoje numa vasta gama de aplicações, podendo-se incluir dentre elas: *processamento de sinais* (cancelamento de ruídos, reconhecimento de voz e caracteres, codificação, compressão de dados, diagnósticos médicos e outros); *controle de sistemas dinâmicos* (manipuladores mecânicos, processos industriais, veículos auto-guiados e outros; *sistemas de decisão* (terapêutica médica, análise financeira e outros); além de numerosas possibilidades descritas na literatura especializada.

Com os conhecimentos obtidos até agora, pode-se considerar que as redes neurais artificiais, futuramente, poderão ter um papel importantíssimo em nossa vida, tanto nas áreas profissionais, no auxílio a tomadas de decisões, quanto nas científicas, proporcionando novas descobertas, principalmente no que diz respeito ao cérebro humano.

Para uma melhor evolução das redes neurais artificiais, deve-se investigar com mais detalhes a capacidade de desenvolver critérios de relevância.

Esta investigação se propõe a ajudar na evolução das redes neurais artificiais em relação à auto-organização.

Nesse sentido, as redes neurais artificiais tornaram-se um assunto interdisciplinar, visto que atraem o interesse de diferentes profissionais, tais como psicólogos, neurofisiologistas, engenheiros, cientistas cognitivos e cientistas da computação, que buscam, cada um em sua área, novas descobertas por meio da computação neural.

Futuramente, pretendemos avançar com as investigações sobre edes Neurais Artificiais, incluindo o *processador quântico*, e criando, se possível, uma metodologia para reconhecimento de classificação das características profissionais de pessoas, sobretudo na FEMA (Fundação Educacional do Município de Assis). Procedimento que poderá ser utilizado, também, na avaliação do perfil de alunos da instituição, podendo assim, talvez, direcioná-los para uma área mais adequada à sua área profissional.

REFERÊNCIAS BIBLIOGRÁFICAS

ASHBY, W. R. Principles of the Self-Organizing System. In: VON FOERSTER, H.; ZOPF JR., G.W., (org.). *Principles of Self-Organization*. Oxford: Pergamon, 1962. pp. 255-278.

_____. *Uma Introdução à Cibernética*. São Paulo: Perspectiva, 1970.

ATLAN, H. *Entre o cristal e a fumaça: ensaio sobre a organização do serviço*. Rio de Janeiro: Jorge Zahar, 1992.

BARLOW, H. B. Unsupervised learning. *Neural Computation*. 1989. vol. 1, pp. 295-311.

BRAGA, A.de P.; CARVALHO, A. C. P. de L. F.; LUDEMIR, T. B. *Fundamentos de Redes Neurais Artificiais*. XI Escola Brasileira de Computação da SBC, 1998.

BRESCIANI FILHO, E.; D'OTTAVIANO, I. M. L. Conceitos Básicos de Sistêmica. In: D'OTTAVIANO, I. M. L; GONZALES, M. E. Q.. (orgs). *Auto-organização: estudos interdisciplinares e História da Ciência*. 2000.Coleção CLE, v.30, cap.12, pp. 283-306.

CAUDILLl, M; BUTLER, C. *Understanding Neural Networks*. Cambridge (EUA): MIT Press, 1992. vols. 1 e 2

D'OTTAVIANO, I. M. L.; GONZALES, M. E. Q.. *Auto-organização: estudos interdisciplinares e História da Ciência*. 2000. Coleção CLE, v.30, cap.12, pp. 283-306.

DEBRUN, M.. A Dinâmica da Auto-organização Primária. In: DEBRUN, M.; GONZALES, M. E. Q.; PESSOA JR., O., (orgs). *Auto-Organização: estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes*. Campinas: Unicamp, Centro de Lógica, Epistemologia e História da Ciência, 1996. Coleção CLE, v.18, cap.2, pp. 25-59.

DEBRUN, M.; GONZALEZ, M. E. Q.; PESSOA JR., O. *Auto-Organização: estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes*. Campinas: Unicamp, Centro de Lógica, Epistemologia e História da Ciência, 1996. Coleção CLE, v.18.

DUPUY, J. P. *Nas Origens das Ciências Cognitivas*. São Paulo: Editora da Universidade Estadual Paulista, 1996.

EPSTEIN, Isaac. *Teoria da Informação*. São Paulo: Ática, 1986.

GONZALES, M. E. Q.. Ação, Causalidade e Ruído nas Redes Neurais Auto-organizadas. In: DEBRUN, M.; GONZALES, M. E. Q.; PESSOA JR., O. (orgs). *Auto-Organização: estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes*. Campinas: Unicamp, Centro de Lógica, Epistemologia e História da Ciência. 1996. Coleção CLE, v.18, cap.10, pp.273-297.

GRISSON, D.; DIEGUES, A. *Biblioteca Universal de Informática*. São Paulo: D'Livros Editora, 1999.

HAYKIN, S. *Redes neurais: princípios e prática*. 2. ed.. Porto Alegre: Bookman, 2001.

HECHT, N. R. *Neurocomputing*. New York: Addison-Wesley Inc., 1990.

HOPFIELD, J. J.. Neural networks and physical systems with emergent collective properties. *Proceedings of the National Academy of Sciences*. USA, 1982. cap. 79, pp. 2554-2558.

Informática e suas aplicações. In: *OMNIA - Revista de Faculdade de Filosofia, Ciência e Letras de Adamantina – FAFIA*. Adamantina (SP): Editora OMNIA, Semestral, 1998. v.1, p.67.

KOHONEN, T.. *An Introduction to Neural Computing*. Finland: Helsinki University of Tecnology, 1988.

_____. *Self-Organization and Associative Memory*. 2. ed. USA: Spring-Verlag, 1989.

_____. The self-organizing map. *Proceedings of the Institute of Electrical and Electronics Engineers*, 1990. vol.78, pp. 1464-1480.

KOVÁCS, Z. L.. *O Cérebro e a sua Mente: uma introdução à neurociência computacional*. São Paulo: Edição Acadêmica, 1997.

_____. *Redes Neurais Artificiais: Fundamentos e Aplicações*. 2. ed.. São Paulo: Edição Acadêmica, 1996.

LIPPMANN, R.. *An Introduction of Computing with Neural Nets*. IEEE Computer Society, abr.1987. v.3, nº 4, pp. 4-22.

LITTLEJOHN, S. W. *Fundamentos Teóricos da Comunicação Humana*. E.U.A.: Humboldt State University, 1982.

MINSKY, M.; PAPERT, S.. *Perceptrons: an introduction to computational geometry*. Massachusetts: MIT Press, 1960.

MARAR, João Fernando. *Polinômios Potências de Sigmóides (PPS): uma nova técnica para aproximação de funções, construção de wavenets e suas aplicações em processamento de imagens e sinais.* 1997. Tese (Doutorado em Informática) – Universidade Federal de Pernambuco. CCEN, Recife.

MENDEL, J. M.; MCLAREN, R. W. Adaptive, Learning, and Pattern Recognition Systems. *Theory and Applications, chapter Reinforcement-learning control and pattern recognition systems.* New York: Academic Press, 1970. pp. 287-318.

NASCIMENTO, C. L.; YONEYAMA, T.. *Inteligência Artificial em controle e automação*. São Paulo: Editora Edgard Blucher Ltda, 2000.

PEREIRA JR., A.; GONZALES, M. E. Q.. Informação, Organização e Linguagem. In: ÉVORA, F. R. R.. *Espaço e Tempo.* Campinas: Unicamp, Centro de Lógica, Epistemologia e História da Ciência, 1995. Coleção CLE, v.15, cap.13, pp.255-290.

RICH, E.; KNIGHT, K.. *Inteligência Artificial*. 2. ed. São Paulo: Makron Books do Brasil, 1993.

SCHUTZER, H. D. N.. O Problema da Mente na Ciência Cognitiva. In: *Documents IEA* . São Paulo: USP, 1992.

SCHUTZER, H. D. N.; PIQUEIRA, J. R. A multilayer perspective of human cognition. In: *Documents IEA*: São Paulo: USP, 1992.

_____. Cognition: determinism, chaos and emergence. *Manuscript accepted for the IEEE System, Man and Cybernetics Conference Proceedings*. 1992.

SHANNON, C.; WEAVER, W.. Teoria Matemática da Comunicação. São Paulo: Difel, 1977.

SILVA, J. J. da. Informação e Auto-Organização. In: DEBRUN, M.; GONZALES, M. E. Q.; PESSOA JR., O. (orgs.). *Auto-Organização: estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes.* Campinas: Unicamp, Centro de Lógica, Epistemologia e História da Ciência. 1996. Coleção CLE, v.18, cap.5, pp.103-128.

cm 1 2 3 4 5 6 7 unesp 1 12 13 14 15 16 17 18

SOARES, A. B. Q. Modelos Cognitivos da Aprendizagem. In: GONZALES, M. E. Q. (org.). *Encontro com as ciências cognitivas.* Marília: Unesp, Faculdade de Filosofia e Ciências, 1997. Capítulo: Linguagem, Semiótica e Cognição, pp.47-62.

TAFNER, M. A. *Revista Cérebro & Mente.* Campinas: Universidade Estadual de Campinas, 1998.

THAGARD, P.. *Mente: introdução à ciência cognitiva.* Porto Alegre: ArtMed, 1998.

UNIVERSITY OF CALIFORNIA. Information & Computer Science eletrônico: banco de dados. Disponível em <http://www.ics.uci.edu/pub/machine-learning-databases/zoo/zoo.names>. Acesso em: 08 set. 2001.

VIANA, F. L.; MARRANGHELLO, N.. *Estudo do Modelo de Kohonen de Redes Neurais.* 1994. (Doutorado) Departamento de Ciências da Computação e Estatística, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto.

VON DER MALSBURG., C. Network self-organization. In: *An Introduction to Neural and Electronic Networks S. F. Zornetzer, JL. Davis, and C. Lau*, San Diego: Academic Press, 1990. pp. 421-432,.