Parte I

Capítulo

6

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Redes Neurais Artificiais

**6.1 Introdução**

As Redes Neurais Artificiais (RNA), também conhecidas como métodos conexionistas, são modelos matemáticos que se assemelham às estruturas biológicas e que têm capacidade computacional adquirida por meio de aprendizado e generalização (Haykin, 2001; Negnevitsky, 2005).

Inicialmente será discutido neste capítulo a representação de conhecimento utilizada pelas Redes Neurais Artificiais, para depois tentarmos analisar a parte referente ao aprendizado. A princípio, é importante salientar que existem diferentes tipos de Redes Neurais Artificiais (**RNA**s) e que cada uma delas tem características próprias em relação a sua representação e aquisição de conhecimento (ou aprendizado) (Osório, 1999).

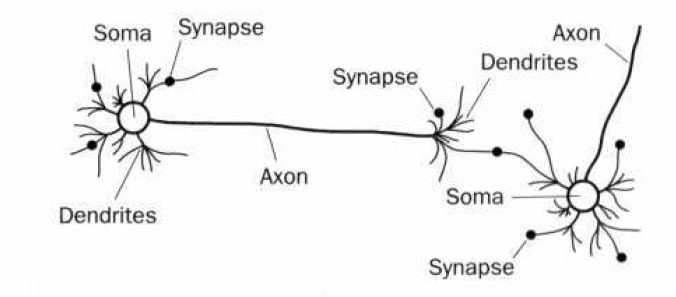
**6.2 Definição**

Segundo Negnevitsky (2005), uma Rede Neural pode ser definida como um modelo de raciocínio baseado no cérebro humano. O cérebro é composto por um conjunto densamente interligado de células nervosas, ou unidades básicas de processamento, chamada de neurônios. O cérebro humano incorpora cerca de 10 bilhões de **neurônios** e 60 trilhões de conexões, **sinapses**, entre eles (Shepherd e Koch, 1990; citado por Negnevitsky, 2005). Por usar múltiplos neurônios simultaneamente, o cérebro pode realizar suas funções muito mais rapidamente do que o computador mais rápido existente hoje.

Embora cada neurônio tenha uma estrutura muito simples, um exército desses elementos possui um tremendo poder de processamento. Um neurônio consiste de um corpo celular, **soma**, um número de fibras chamadas de **dendritos**, e uma fibra longa chamada de **axônio**. Enquanto os dendritos formam uma rede em torno do corpo celular (soma), o axônio se estende para os dendritos e soma de outros neurônios. A Figura 6.1 é um esquema de uma rede neural (Negnevistsky, 2005).

Os sinais são propagados de um neurônio para outro através de complexas reações eletroquímicas. As substancias químicas liberadas das sinapses causam mudanças no potencial elétrico do corpo celular. Quando o potencial atinge seu limite, um pulso elétrico, é enviado através do axônio. O pulso se espalha e eventualmente atinge as sinapses, causando o aumento ou diminuição de seu potencial. No entanto, a descoberta mais interessante é que a rede neural exibe plasticidade (Negnevistky, 2005).

**Figura 6.1** Rede Neural Biológica. **Fonte**: Artificial intelligence: a guide to intelligence systems/Michael Negnevitsky, 2005).

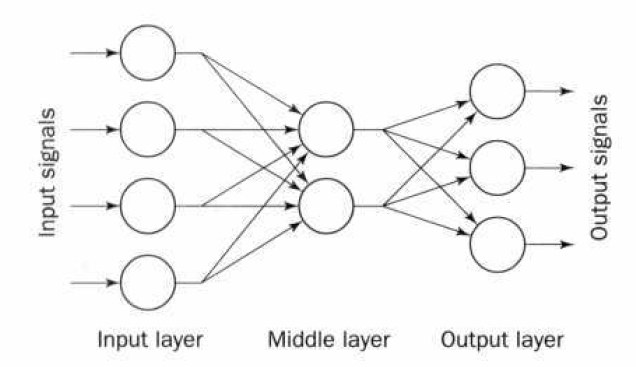


Um neurônio em “desenvolvimento” é sinônimo de um cérebro plástico: a plasticidade permite que o sistema nervoso em desenvolvimento se adapte ao meio ambiente. Assim com a plasticidade parece ser essencial para o funcionamento dos neurônios como unidades de processamento de informações do cérebro humano, também ela o é em relação às redes neurais construídas com neurônios artificiais (Haykin, 2001).

Portanto, de acordo com Negnevitsky (2005), pode-se definir as redes conexionistas ou Rede Neuronal Artificial (RNAs), como sendo formadas por um conjunto de unidades elementares de processamento de informações fortemente conectadas, que denomina-se neurônio artificial. Uma RNA é constituída por um grafo orientado e ponderado. Os nós do grafo são autômatos simples, os chamados neurônios artificiais, que formam através de suas conexões um autômato mais complexo, a rede neural. A Figura 6.2 ilustra uma rede neural artificial.

Segundo Negnevitsky (2005), cada unidade da rede é dotada de um estado interno, que geralmente é denominado de estado de ativação. As unidades podem propagar seu estado de ativação para outras unidades do grafo, passando pelos arcos ponderados, que é chamado de conexões, ligações sinápticas, ou simplesmente de pesos sinápticos. A regra que determina a ativação de um neurônio em função da influência transmitidas de suas entradas, ponderadas pelos seus respectivos pesos, se chama regra de ativação ou função de ativação. E a Tabela 6.1 faz uma analogia entre uma rede biológica e uma rede conexionista.

**Figura 6.2** Arquitetura de uma típica Rede Neural Artificial. **Fonte**: Artificial intelligence: a guide to intelligence systems/Michael Negnevitsky, 2005).



**Tabela 6.1** Analogia entre Rede Neural biológica e Artificial

|  |  |
| --- | --- |
| Rede neural biológica | Rede neural artificial |
| Soma | Neuron |
| Dendrite | Input |
| Axon | Output |
| Synapse | weight |

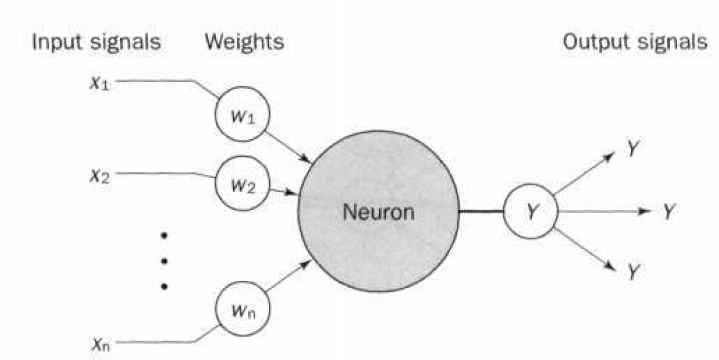
O processamento da informação em RNAs é realizado por meio de estruturas neurais artificiais em que o armazenamento e o processamento da informação são realizados de maneira paralela e distribuída por elementos processadores relativamente simples. Cada elemento processador corresponde a um neurônio artificial (Negnevistsky, 2005).

Segundo Negnevistsky (2005) para se construir uma RNA, deve-se decidir primeiro como os neurônios serão usados e como eles serão conectados na rede. Em outras palavras, deve-se escolher a arquitetura da rede. Então, decide-se que algoritmo de aprendizado usar. E finalmente, treina-se a rede, isto é, inicializa-se os pesos da rede e atualiza-se os pesos para o conjunto de treinamento.

**6.3 O Neurônio**

Um neurônio recebe vários sinais através de seus links de entrada (input), calcula um novo nível de ativação, e envia-o como sinal de saída através dos links de saída. O sinal de entrada pode ser uma linha de dados ou a saída de outro neurônio. O sinal de saída pode ser a solução para o problema ou uma entrada para outro neurônio. A Figura 6.3 mostra um típico neurônio.

**Figura 6.3** Diagrama de um Neurônio. **Fonte**: Artificial intelligence: a guide to intelligence systems/Michael Negnevitsky, 2005).



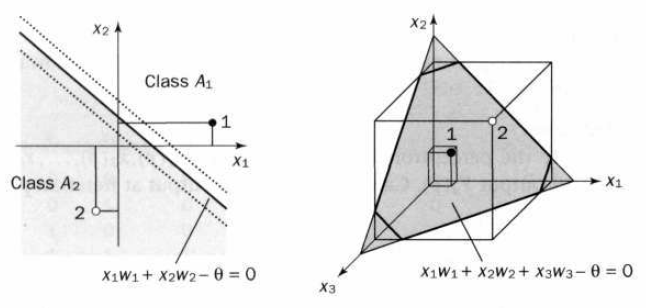
As mudanças realizadas nos valores dos pesos sinápticos ou na estrutura de interconexão das unidades de uma rede, são responsáveis pelas alterações no comportamento de ativação desta rede. Estas alterações nas conexões e na estrutura da rede é o que nos permite realizar o aprendizado de um novo comportamento. Desta maneira, pode-se modificar o estado de ativação da saída da rede em resposta a uma certa configuração de entradas. Dessa forma, a rede é capaz de estabelecer associações de entrada-saída (estímulos e respostas) a fim de se adaptar a uma situação proposta. O método utilizado para modificar o comportamento de uma rede é denominado de *regra de aprendizado* (Osório, 1999).

O primeiro modelo matemático do neurônio foi o modelo de proposto por McCulloch e Pitts, em 1943. Mais tarde Rosenblatt (1957) criou o modelo do perceptron. O modelo consiste de um modelador linear seguido de limitador. Um perceptron modela um neurônio tomando uma soma ponderada de suas entradas e compara essa soma com o limiar, o que produz uma saída +1 se suas entradas são positivas e -1 se elas são negativas. O objetivo do *perceptron* é classificar as entradas, ou em outras palavras, aplicar externamente x1, x2, ..., xn estímulos entre umas das classes, ditas A1 e A2. Assim, no caso de um *perceptron* elementar, o espaço n-dimensional é dividido por um hiperplano em duas regiões. O hiperplano é definido por função linearmente separáveis.

**(6.1)**

Para o caso de duas entradas x1 e x2, o limite de decisão leva a forma de uma linha reta como ilustrada na Figura 6.4(a). O ponto, o qual se encontra acima da linha limite, pertence a classe A1; e o ponto 2, o qual se encontra abaixo da linha limite, pertence à classe A2.

**Figura 6.4** Separabilidade linear no perceptrons: (a) duas entradas; (b) três entradas. **Fonte**: Artificial intelligence: a guide to intelligence systems/Michael Negnevitsky, 2005).



1. (b)

O neurônio de McCulloch e Pitts, usa a seguinte transferência ou função de ativação:

**(6.2)**

Onde X é a soma ponderada das entradas do neurônio, e xi é o valor da entrada **i**, wi é o peso da entrada **i**, **n** é o número de neurônios de entrada e **y** é o número de neurônios de saída.

Este tipo de função de ativação é chamada de função sinal (*sign function*).

Assim, a saída do neurônio para uma função sinal, pode ser representado como

**(6.3)**

Segundo Negnevistsky (2005) muitas funções de ativação têm sido testadas, mais somente umas poucas tem tido aplicações práticas. Quatro delas são a função, step, sign, linear e sigmoide.

* As funções de ativação step e sign – também chamadas de ***hard limit functions***, são frequentemente usadas em neurônios para tomada de decisões para tarefas classificação e reconhecimento de padrões.
* A função sigmoide – transforma as entradas, que pode ter qualquer valor entre mais ou menos infinito, em um valor razoável na faixa de 0 e 1.
* A função linear – neurônios com função de ativação linear são frequentemente usadas para aproximações lineares.

**6.4 Classificação e Propriedades**

A grande variedade de modelos existentes nos leva a um estudo ou análise das principais propriedades das redes neurais, que nos permita compreender melhor as vantagens e/ou inconveniências da escolha de um modelo em detrimento de outro. Considere-se que para essa análise sejam avaliados um grupo de atributos tais como: tipo de aprendizado, arquitetura de interconexões, forma interna de representação das informações, tipo de aplicação da rede, etc (Rosa, 1999).

**6.4.1 Aprendizado RNA**

Uma RNA aprende geralmente de forma gradual, onde os pesos são modificados várias vezes, paulatinamente, seguindo-se uma regra de aprendizado que estabelece a forma como estes pesos são alterados. Utiliza-se no aprendizado um conjunto de dados de aprendizado disponível (base de exemplos). Cada iteração deste processo gradativo de adaptação dos pesos de uma RNA, é chamada de época de aprendizado. Os métodos de aprendizado neural podem ser divididos em três grandes classe, segundo a grau de controle dado ao usuário (Rosa, 1999; Rezende, 2003):

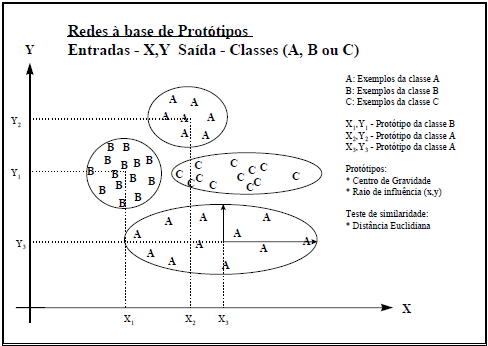
* **Aprendizado supervisionado**: o usuário dispõe de um comportamento de referência preciso que ele deseja ensinar a rede. Sendo assim, a rede deve ser capaz de medir a diferença entre seu comportamento atual e o comportamento de referência, e então corrigir os pesos de maneira a reduzir este erro (desvio de comportamento em relação aos exemplos de referência).
* **Aprendizado semi-supervisionado**: o usuário possui apenas indicações imprecisas (por exemplo: sucesso/insucesso da rede) sobre o comportamento final desejado. As técnicas de aprendizado semi-supervisionado são chamadas também de aprendizado por reforço (reinforcement learning) [Sutton 98].
* **Aprendizado não-supervisionado**: os pesos da rede são modificados em função de critérios internos, tais como, por exemplo, a repetição de padrões de ativação em paralelo de vários neurônios. O comportamento resultante deste tipo de aprendizado é usualmente comparado com técnicas de análise de dados empregadas na estatística (*clustering*).
* **Aprendizado instantâneo**: o conjunto de dados de aprendizado é analisado uma única vez e com isto o conjunto de pesos da rede é determinado de maneira imediata em uma única passagem da base de exemplos. Este modo de aprendizado também é conhecido como: *one single epoch learning* / *one shot learning*.
* **Aprendizado por pacotes:** o conjunto de dados de aprendizado é apresentado à rede várias vezes, de modo que possamos otimizar a resposta da rede, reduzindo os erros da rede e minimizando o erro obtido na saída desta. Este modo de aprendizado é caracterizado por trabalhar com uma alteração dos pesos para cada época, ou seja, para cada passagem completa de todos os exemplos base de aprendizado. O algoritmo de aprendizado deve reduzir pouco à pouco o erro de saída, o que é feito ao final de cada passagem (análise) da base de exemplos de aprendizado.
* **Aprendizado continuo:** o algoritmo de aprendizado leva em consideração continuamente os exemplos que lhe são repassados. Se o conjunto de dados é bem delimitado, chamamos este método de aprendizado on-line, e caso o conjunto de dados possa ir aumentando (sendo adicionados novos exemplos no decorrer do tempo), então chamamos este método de aprendizado incremental. O aprendizado on-line se opõe ao aprendizado por pacotes, pois ao contrário deste, para cada novo exemplo analisado já se realiza uma adaptação dos pesos da rede, com o objetivo de convergir na direção da solução do problema. O principal problema do aprendizado contínuo é a dificuldade de achar um bom compromisso entre a plasticidade e a estabilidade da rede. Uma rede com uma grande facilidade de adaptação pode “esquecer” rapidamente os conhecimentos anteriormente adquiridos e uma rede com uma grande estabilidade pode ser incapaz de incorporar novos conhecimentos.
* **Aprendizado ativo:** este modo de aprendizado assume que o algoritmo de adaptação da rede pode passar de uma posição passiva (apenas recebendo os dados do jeito como lhe são passados), para uma posição ativa. Sendo assim, assumimos que este algoritmo poderá vir a intervir sobre a forma como os dados lhe são repassados. Neste caso, a rede pode intervir e determinar assim quais dados que serão considerados e/ou desconsiderados, além também de determinar a ordem em que estes dados deverão ser considerados. A rede pode também vir a solicitar novos dados que julgue necessários para o bom aprendizado do problema proposto.

**6.4.2 Tipos de Unidades**

Segundo Rosa (1999), as unidades de uma rede – os neurônios artificiais – podem ser de diferentes tipos, de acordo com a função interna utilizada para calcular o seu estado de ativação, ou seja, qual a função de ativação utilizada linear, gaussiana, sigmoide, assimétrica, etc. Um outro elemento que pode diferenciar uma unidade, diz respeito a forma como os neurônio armazenam as informações: unidades baseadas em protótipos, unidades do tipo Perceptron.

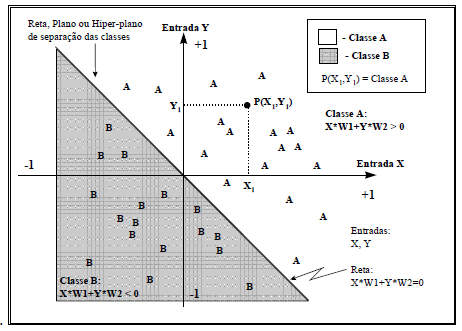
* **Redes à base de protótipos**: este tipo de rede utiliza neurônios que servem para representar protótipos dos exemplos aprendidos – as unidades tem uma representação interna que agrupa as características comuns e típicas de um grupo de exemplos (Orsier 95). As redes baseadas em protótipos tem normalmente um aprendizado não supervisionado (com um ou mais protótipos associados à cada classe). Uma das vantagens deste tipo de redes é a possibilidade de fazer um aprendizado contínuo e incremental, uma vez que não é muito difícil de conceber um algoritmo capaz de aumentar a rede neural através da adição de novos protótipos. Os protótipos são também denominados de clusters, onde apresentamos um exemplo de rede a base de protótipos na Figura 6.5.

**Figura 6.5** Protótipo de uma rede neural com duas entradas. **Fonte**: Osório, Fernando. Redes Neurais – Aprendizado Artificial. Forum de I.A. “99 – pg.13”.



* **Redes à base de Perceptrons**: as unidades do tipo “*Perceptron*” foram criadas por Frank Rosenblatt em 1950. Este é um dos modelos de neurônios mais utilizados na atualidade. Ele é a base de diversos tipos de RNA com aprendizado supervisionado utilizando uma adaptação por correção de erros (usualmente baseada na descida da superfície de erro usando o gradiente). O modelo do Perceptron de múltiplas camadas (**MLP** – *Multi-Layer Perceptron*) tornou-se muito conhecido e aplicado, sendo na maior parte das vezes associado a regra de aprendizado do *Back-Propagation* (Jodoin 94, Widrow 90, Rumelhart 86). A Figura 6.6 apresenta um esquema da representação de conhecimentos nas redes baseadas em *Perceptrons*, e como este tipo de redes é capaz de classificar padrões, gerando planos (ou hiperplanos) de divisão do espaço em que se situam os exemplos (Negnevitsky, 2005).

**Figura 6.6** Separação de classes (classificação) através do uso de um Perceptron. **Fonte**: Osório, Fernando. Redes Neurais – Aprendizado Artificial. Forum de I.A. “99 – pg.13”.

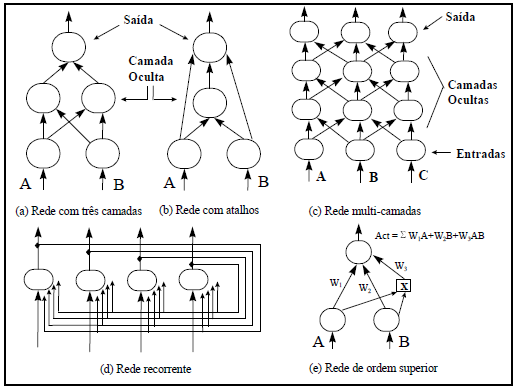


**6.4.3 Tipos de Arquiteturas de Conexões de Redes**

As unidades de uma rede neural podem se conectar de diferentes modos, resultando em diferentes arquiteturas de interconexão de neurônios. A Figura 6.7 apresenta alguns exemplos de possíveis maneiras de conectar os componentes de uma RNA. As arquiteturas mais importantes são (Osório, 1999):

* **Redes com uma única camada**: as unidades estão todas em um mesmo nível. Neste tipo de arquitetura, as unidades são conectadas diretamente às entradas externas e estas unidades servem também de saídas finais da rede. As redes de uma única camada possuem normalmente conexões laterais (entre os neurônios de uma mesma camada). Um exemplo deste tipo de arquitetura de redes são as redes do tipo “*Self-Organizing Feature Maps*” (Kohonen,1987).
* **Redes com camadas unidirecional**: as unidades são organizadas em vários níveis bem definidos, que são chamados de camadas ou *layers*. Cada unidade de uma camada recebe suas entradas vindas à partir de uma camada precedente, e envia seus sinais de saídas em direção a camada seguinte. Estas redes são conhecidas como redes *feed-forward*. A Figura 6.7(a) mostra um exemplo de uma rede de três camadas unidirecionais. Esta arquitetura de três camadas (entrada, camada oculta e saída) é muito usada em aplicações práticas das redes neurais.
* **Redes recorrentes**: as redes recorrentes podem ter uma ou mais camadas, mas a sua particularidade reside no fato de que temos conexões que partem da saída de uma unidade em direção a uma outra unidade da mesma camada ou de uma camada anterior à esta. Este tipo de conexões permitem a criação de modelos que levam em consideração aspectos temporais e comportamentos dinâmicos, onde a saída de uma unidade depende de seu estado em um tempo anterior. Os laços internos ao mesmo tempo que dão características interessantes de memória e temporalidade as redes, tornam este tipo de redes muito instáveis, o que nos obriga a usar algoritmos específicos (e usualmente mais complexos) para o aprendizado destas redes. Um tipo particular de redes recorrentes são as redes totalmente conectadas, e um exemplo de modelo recorrente de uma única camada e totalmente conectado são as redes de Hopfield, representadas na figura 6.7(d).
* **Redes de ordem superior:** as unidades deste tipo de rede permitem a conexão direta entre duas ou mais de suas entradas, antes mesmo de aplicar a função de cálculo da ativação da unidade (Fiesler, 1994a). Este tipo de rede serve para modelar “sinapses de modulação”, ou seja, quando uma entrada pode modular (agir sobre) o sinal que vem de uma outra entrada. Um modelo particular de rede de ordem superior são as redes tipo *Sigma-Pi* que foram apresentadas no livro PDP – *Parallel Distributed Processing* (Rumelhart 86), e que são representadas na figura 6.7(e).

**Figura 6.7** Arquiteturas de interconexão de neurônios em redes. **Fonte:** Osório, Fernando. Redes Neurais – Aprendizado Artificial. Forum de I.A. “99 – pg.13”.



A arquitetura de uma rede também pode ser classificada de acordo com a evolução desta no decorrer de sua utilização e desenvolvimento do aprendizado. Em relação a este critério pode-se ter os seguintes tipos (Osório, 99):

* **Redes com estruturas estáticas**: a rede tem a sua estrutura definida antes do início do aprendizado. A quantidade de neurônios, assim como a sua estrutura de interconexões, não sofrem alterações durante a adaptação da rede. As únicas mudanças se realizam à nível dos pesos sinápticos, que são modificados durante o processo de aprendizado (Osório, 98).
* **Redes com estruturas dinâmicas**: as redes que possuem uma estrutura dinâmica são redes onde o número de unidades e conexões pode variar no decorrer do tempo. Estas redes são também chamadas de *ontogênicas* (Fiesler, 1994). As modificações na estrutura da rede podem ser do tipo generativo (incremental) ou do tipo destrutivo (redutor por eliminação/simplificação). A escolha entre estes dois tipos de métodos é bastante polêmica: devemos começar com uma rede pequena e ir aumentando ela, ou devemos começar com uma rede bastante grande e ir reduzindo o seu tamanho posteriormente? Alguns autores defendem a ideia de uma criação construtiva de conhecimentos (Elman 1993, Osório 1999).

**6.5 Tipos de Aplicações para Redes Neurais**

De acordo com diversos autores, as RNAs podem ser aplicadas a diversos tipos de tarefas, tais como: o reconhecimento de padrões (e.g. reconhecimento de faces humanas), a classificação (e.g. reconhecimento de caracteres – **OCR**), a transformação de dados (e.g. compressão de informações), a predição (e.g. previsão de séries temporais, como as cotações da bolsa de valores, ou o uso para diagnósticos médicos), o controle de processos e a aproximações de funções (e.g. aplicações para área de robótica). Todas essas tarefas podem ser agrupadas em dois grandes grupos (Osório, 1999): **Redes para aproximações de funções**, **Redes para classificação de padrões**.

**6.6 Vantagens das Redes Neurais**

De acordo com Osório (1999), as redes conexionistas, em particular aquelas comumente aplicadas na construção de sistemas inteligentes, apresentam as seguintes vantagens:

* **Conhecimento empírico**: em geral as redes aprendem mais fácil do que outros métodos de aquisição de conhecimento, pois o aprendizado acontece a partir de exemplos de maneira simples e permite um aquisição de conhecimento de forma automática.
* **Degradação progressiva**: apesar das redes serem menos sensíveis as perturbações, do que os sistemas simbólicos. As respostas dadas por uma rede se degrada progressivamente na presença de perturbações e distorções dos dados de entrada.
* **Manipulação de dados quantitativos:** as redes trabalham com a representação numérica dos conhecimentos e, isso implica que as redes são melhor adaptadas para a manipulação de dados quantitativos (valores contínuos). Isso pode ser considerado uma vantagem, uma vez que grande parte dos problemas do mundo real, manipulam valores contínuos.
* **Paralelismos em larga escala:** as redes neurais são compostas de um conjunto de unidade de processamento de informações que podem trabalhar em paralelo. Apesar da maioria das implementações de RNAs serem feitas através de simulações em máquinas sequenciais, é possível de se implementar (softwares e hardwares) que possam explorar esta possibilidade de ativação simultânea das unidades de uma rede. A maior parte das implementações de redes neurais simuladas em máquinas sequenciais pode ser facilmente adaptada em uma versão paralela deste sistema.

**6.7 Inconvenientes das Redes Neurais**

As redes apresentam alguns inconvenientes, do mesmo modo que outros tipos de métodos de aprendizado. As redes apresentam os seguintes inconvenientes (Osório, 1999):

* **Arquitetura e parâmetros**: a evolução do processo de aprendizado é bastante influenciado por estes dois parâmetros. Como não existe métodos totalmente automatizados para escolha correta da arquitetura para um problema, fica muito difícil de se encontrar uma boa topologia para a rede, bem como, bons parâmetros de regulagem para o algoritmo de aprendizado. O sucesso da rede depende bastante desses dois elementos, que variam muito de um problema para outro.
* **Inicialização e codificação**: uma má escolha dos pesos iniciais da rede, do método de codificação dos dados de entrada, ou mesmo da ordem de apresentação destes, pode levar ao bloqueio do processo de aprendizado, ou pode dificultar o processo de convergência da rede na direção de uma boa solução. Uma vez que, os algoritmos de aprendizado conexionistas são em geral muito dependentes do estado inicial da rede e da codificação dos dados da base de aprendizado.
* **Caixa preta**: as redes conexionistas são “caixas preta” onde os conhecimentos ficam codificados de tal forma que estes são inteligíveis para o utilizador ou até mesmo para um especialista. Isto pelo fato, de que, os conhecimentos adquiridos por uma rede estão codificados no conjunto de valores dos pesos sinápticos, e também pela maneira pela a qual as unidades se conectam.
* **Conhecimento teórico**: como as árvores de decisão, as redes neurais são orientadas para a aquisição de conhecimentos empíricos (baseados em exemplos). Um modo simplista de se aproveitar algum conhecimento teórico pré-existente, consiste em se converter regras em exemplos (“protótipos” representativos destas regras). Entretanto, este tipo de método não nos garante que a rede será capaz de aprender corretamente estes exemplos, sendo assim, não podemos garantir que ao final do aprendizado todos os conhecimentos teóricos disponíveis estarão bem representados internamente na rede.

Nessa sessão foram apresentadas algumas das vantagens e desvantagens das redes conexionistas, sem a pretensão de cobrir exaustivamente todas elas, mas apenas para se ter uma ideia das principais características deste tipo de sistema.

**6.8 Conclusões**

Neste capítulo apresentou-se uma visão geral sobre os sistemas de I.A. e a necessidade do aprendizado para que um sistema inteligente possa ser considerado como tal. Dando ênfase ao aprendizado neural como sendo uma forma de aquisição de conhecimentos, que dadas as suas peculiaridades, possui um interesse particular na área de inteligência Artificial.

Considerando-se suas principais características: a representação de conhecimentos, o paralelismo inerente as unidades da rede, a sua capacidade de adaptação, entre outros aspectos. No entanto, observa-se que as redes neurais possuem ainda alguns pontos fracos a serem estudados, principalmente no que diz respeito a explicitação dos conhecimentos adquiridos e na dificuldade de convergência em relação a uma solução ótima.