

Redes Neuronais Artificiais

Capítulo 6:

Rocha, M., Cortez, P. e Neves, J. (2008). Análise Inteligente de Dados – Algoritmos e Implementação em Java, FCA.

Introdução

- O **cérebro humano** é uma estrutura altamente não linear e paralela
- Consegue **aprender pela experiência** e é capaz de executar tarefas complexas que os mais poderosos computadores atuais ainda são incapazes de realizar
- Por esse motivo, serviu de inspiração para o desenvolvimento das **Redes Neuronais Artificiais** (RNA), também conhecidas como sistemas connexionistas

Introdução

- As **RNA** modernas surgiram a partir do trabalho de *McCulloch* e *Pitts* (1943), que descrevia um cálculo lógico de RNA, unindo os estudos neurobiológicos com a lógica
- O relançamento do paradigma nestas últimas décadas deve-se, em grande parte, à re-invenção do **algoritmo de retropropagação** (*backpropagation*), através do trabalho de *Rumelhart*, *Hinton* e *Williams* (1986)

Introdução

- Desde então, o crescimento deste campo de investigação tem sido impressionante, tendo sido aplicado nas mais diversas áreas, como por exemplo:
 - **Medicina**
 - No diagnóstico de doenças e previsão do tempo de recuperação de pacientes
 - **Biologia**
 - Na interpretação de sequências do genoma humano
 - **Engenharia**
 - Na deteção de falhas em estruturas de pontes

Introdução

– Economia

- Na previsão do valor de ações na bolsa

– Robótica

- Na visão por computador e condução autónoma

Introdução

- Uma **RNA** é constituída por um conjunto de unidades básicas de processamento, designadas de **neurónios** ou **nodos**
- Estes encontram-se fortemente interligados por **conexões** ou **sinapses**
- Numa **RNA** cada conexão une 2 neurónios, com uma determinada intensidade, dada pelo **peso da conexão**, um valor numérico

Introdução

- O **conhecimento** é adquirido a partir de um ambiente (dados), através de um processo de aprendizagem, sendo armazenado nas conexões entre os nodos
- Este processo, designado por **treino da RNA**, é concretizado computacionalmente por um **algoritmo de treino**
- Após o treino, a RNA é capaz de responder, num curto espaço de tempo, a novas situações

Introdução

- Do ponto de vista da Estatística, **as RNA são modelos funcionais**, revelando-se como ferramentas interessantes para a **análise de dados**, uma vez que possuem uma elevada flexibilidade
- Conforme o tipo de estrutura interna que é escolhida, **as RNA podem modelar funções** lineares simples, funções lineares contínuas e até mesmo complexas funções descontínuas

Introdução

- O treino da RNA passa por fornecer um conjunto de casos de treino, onde cada instância é composta por um vetor de entradas e um outro de saídas
- Uma iteração do algoritmo de treino é composta por ajustamentos iterativos dos pesos das conexões para todos os casos de treino
- A aprendizagem é então conseguida quando, após várias iterações, o erro é reduzido para valores aceitáveis

Introdução

- A partir dos anos 80 foram desenvolvidos diversos tipos de RNA, devido à elevada atenção que estas criaram na comunidade científica, como as
 - *Radial Basis Functions*, ou as
 - *Probabilistic Neural Networks*
- Aqui iremos utilizar as redes do tipo **Perceptrão Multicamada** (*Multilayer Perceptrons – MLP*)
 - *Bishop, 1995; Haykin, 1999.*

Introdução

- Nas redes do tipo **Perceptrão Multicamada** as conexões são unidirecionais, sendo os nodos organizados por camadas
- Estas RNA são as mais populares e as mais utilizadas, devido ao **sucesso do algoritmo de retropropagação**
- Além disso, existe uma prova matemática de que são **aproximadores universais de problemas**, ou seja, em teoria possuem a capacidade de modelar uma qualquer função, **por mais complexa que seja**

Benefícios e Limitações

- As RNA apresentam características poderosas e únicas, tais como (Patterson, 1996; Haykin, 1999):
 - **Processamento maciçamente paralelo**
 - Permitindo que tarefas complexas sejam realizadas num curto espaço de tempo
 - **Aprendizagem e generalização**
 - Conseguindo responder de modo adequado a novas situações com base em experiências passadas
 - **Não linearidade**
 - Atendendo a que muitos dos problemas reais a equacionar e resolver são de natureza não linear

Benefícios e Limitações

- As RNA apresentam características poderosas e únicas, tais como (Patterson, 1996; Haykin, 1999):
 - Adaptabilidade
 - Podendo adaptar a sua topologia de acordo com mudanças do ambiente
 - Robustez e degradação suave
 - Permitindo processar ruído, atributos irrelevantes e informação incompleta de forma eficiente, assim como sendo capazes de manter o seu desempenho quando há desativação de algumas das suas conexões e/ou nodos
 - Flexibilidade
 - Com um grande domínio de aplicabilidade e usabilidade, podem ser utilizadas como “caixas negras”, não sendo necessário um conhecimento explícito acerca da função a ser aprendida

Benefícios e Limitações

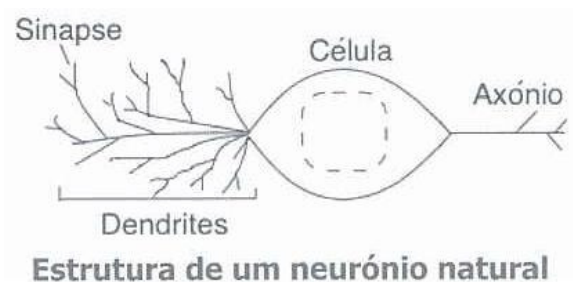
- Isto não quer dizer que as RNA conseguem dar resposta a qualquer problema
- Muitas vezes, necessitam de ser integradas com outros sistemas para a resolução de problemas
- Ainda se está muito longe de atingir uma arquitetura que simule o cérebro humano

Inspiração Biológica

- As RNA foram inspiradas pelo estudo do funcionamento do sistema nervoso central dos seres humanos

Inspiração Biológica

- O seu órgão principal, o cérebro, possui cerca de **10 bilhões de células nervosas (neurónios)**, organizados numa estrutura fortemente interligada por cerca de **60 trilhões de sinapses**, onde um único neurónio pode estar ligado a centenas ou milhares de outros neurónios (*Haykin, 1999*)



Inspiração Biológica

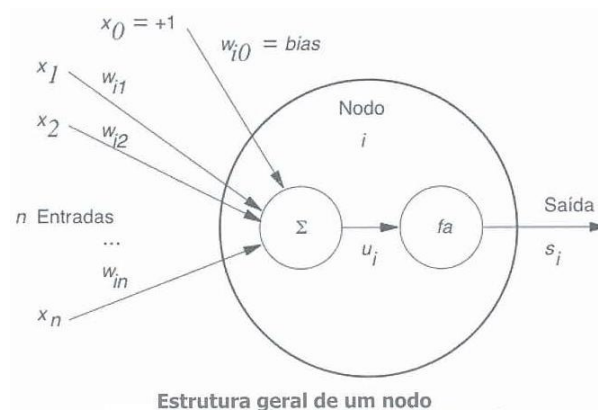
- O modo de processamento da informação no cérebro é eminentemente paralelo, o que leva a que, apesar da **frequência de disparo de um neurónio ser muito inferior à de uma porta lógica**, o cérebro humano seja capaz de efetuar tarefas complexas ainda não ao alcance dos computadores modernos

Inspiração Biológica

- As RNA procuram modelar, de alguma forma, o processamento da informação ao nível do sistema nervoso
- Ainda assim, muitas das características mais complexas dos sistemas biológicos são simplificadas ou ignoradas (*Mitchell*, 1997)
- Por essa razão, **as RNA são abordadas como uma técnica computacional de resolução de problemas** e não como modelos de funcionamento do cérebro humano

Neurónio Artificial

- A estrutura computacional de uma RNA baseia-se no funcionamento dos neurónios, que constituem as suas unidades básicas de funcionamento
- A estrutura de um nodo é relativamente simples

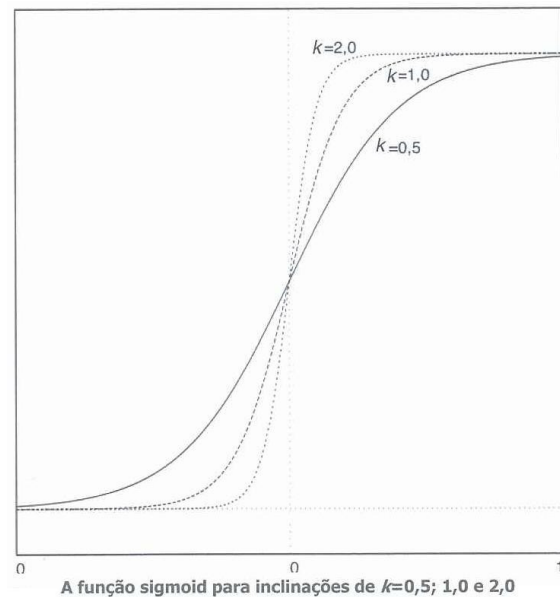


Neurónio Artificial

- As suas principais componentes são:
 - Um conjunto de conexões de entrada, cada uma caracterizada por um peso w_{ij} entre os neurónios j (origem) e i (destino)
 - Assim, a saída do neurónio j (x_j) é multiplicada pelo peso da conexão entre j e i (w_{ij}), sendo assim calculado o valor à entrada do neurónio i
 - A partir das suas entradas, este calcula um único valor, dito de ativação $u_i = \sum x_k w_{ik}$
 - Pode ainda existir, para cada nodo, uma conexão denominada de bias, cuja entrada toma sempre o valor +1, e que adiciona à ativação o valor do seu peso (w_{i0})

Neurónio Artificial

- As suas principais componentes são:
 - Uma **função de ativação (fa)**, que calcula o valor de saída (s_i) do neurónio i , a partir do valor de ativação
 - Esta função é tipicamente não linear, sendo a mais utilizada a função logística ou **sigmoid**, ou seja:
- $$s_i = 1 / (1 + \exp(-ku_i))$$



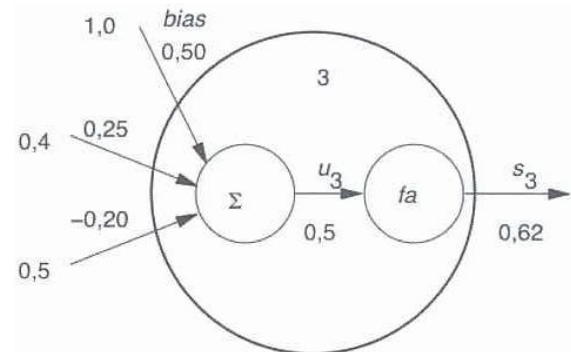
Neurónio Artificial

- Em geral, o valor de inclinação (k) é fixado no valor de 1.0
- A variação do valor da inclinação permite obter funções com diferentes declives
- No limite, quando k se aproxima do infinito, a função tende para uma função **degrau**
- A derivada desta função (importante para a fase de treino) toma a forma de:

$$s'_i = k s_i (1 - s_i)$$
- Outras funções usadas são a tangente hiperbólica (**tanh**), ou a função identidade, ou seja, $s_i = u_i$

Neurónio Artificial

- A título ilustrativo, a figura mostra como se ativa o nodo 3 de uma RNA, com
 - Função de ativação logística padrão ($k=1.0$)
 - Entradas 0,4 e 0,5



Exemplo de ativação de um nodo

sendo que

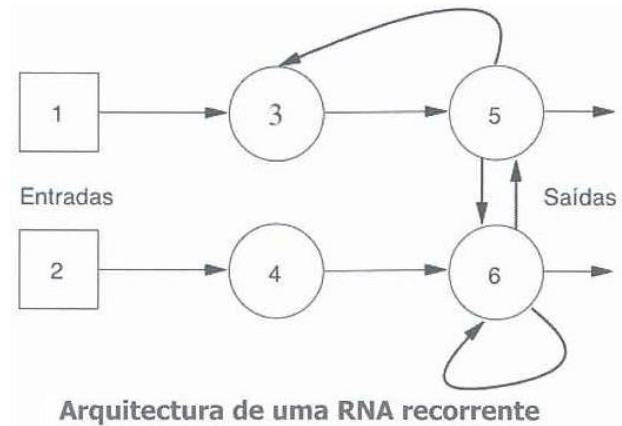
- $u_3 = 1,0 \times 0,5 + 0,4 \times 0,25 + 0,5 \times (-0,20) = 0,5$
- $s_3 = 1 / (1 + \exp(-0,5)) = 0,62$

Topologia

- O modo como os neurónios se interligam numa RNA é designado por **topologia** ou **arquitetura**, podendo ser encarada como um grafo
- As RNA diferem na sua topologia de acordo com o paradigma de aprendizagem que implementam
- Na Aprendizagem Supervisionada (AS) as RNA organizam-se em conjuntos de neurónios de entrada, de saída e intermédios
- Neste caso, os 2 principais tipos de **RNA** denominam-se de **Unidireccionais** ou Perceptrão Multicamada (*Feedforward* ou *Multilayer Perceptrons*) e de **Recorrentes** (*Patterson, 1996*)

Topologia

- A diferença entre os 2 paradigmas diz respeito à possibilidade de existência de ciclos
- De facto, as **redes recorrentes** permitem a ocorrência de ciclos na sua topologia, o que implica que a saída de um dado neurónio pode influenciar a sua própria entrada, criando-se **circuitos fechados**

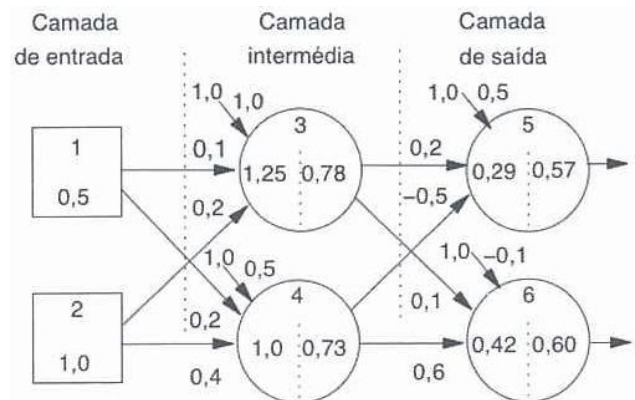


Topologia

- Neste caso, os valores à saída dos neurónios não são função apenas dos valores dos neurónios que os alimentam e dos pesos das conexões, mas também de uma **dimensão temporal**

Topologia

- Por seu lado, as **redes unidireccionais** não permitem a existência de ciclos, o que facilita o cálculo da saída dos neurónios
- No exemplo da figura, primeiro computa-se a saída dos nodos 3 e 4, depois, ativam-se os nodos 5 e 6



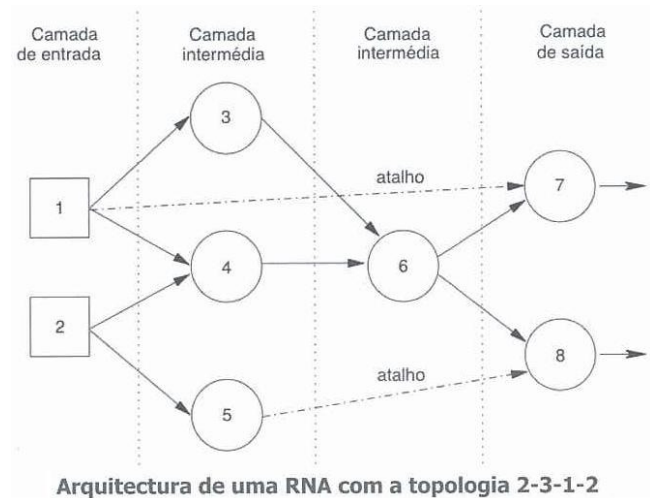
Arquitetura de uma RNA unidirecional com uma camada intermédia, funções de ativação logísticas e conexões de bias

Topologia

- Neste tipo de RNA é possível dividir os neurónios em camadas:
 - Uma **camada de entrada**, cujos valores de ativação são fixados externamente
 - Uma **camada de saída**, cujos valores de ativação representam saídas da RNA
 - E por 0 ou mais **camadas intermédias**

Topologia

- É comum representar esta RNA de uma forma abreviada pela forma $E - I_1 - \dots - I_n - S$
 - E : número de neurónios de entrada
 - S : número de neurónios de saída
 - I_i : cardinalidade da camada intermédia i



Topologia

- Quando existem todas as ligações possíveis entre neurónios de camadas consecutivas, a RNA diz-se **completamente interligada**
- Por vezes existem ligações diretas entre as entradas e os neurónios de saída, que saltam camadas, designando-se por **atalhos**