Redes Neurais Artificiais

Introdução

INFORMAÇÃO,

TECNOLOGIA

& INOVAÇÃO

Introdução

- **Rede neural**: máquina desenvolvida para modelar a maneira com que o cérebro realiza uma tarefa. Neurônios artificiais conectados que passam por um processo de aprendizado .

- **Definição**: sistema de processamento massivamente paralelo e distribuído, construído com unidades de processamento simples, que tem uma propensão natural de armazenar conhecimento por meio de experiência e torná-lo disponível para uso.

Introdução

Uma rede neural se assemelha a um cérebro em dois aspectos:

- Conhecimento é adquirido do ambiente pela rede por meio de um processo de aprendizado.

- As forças das conexões entre os neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

Introdução

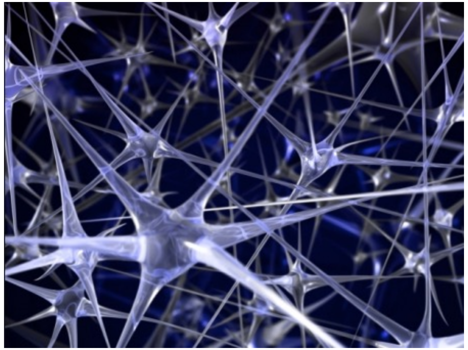
Inspiração biológica:

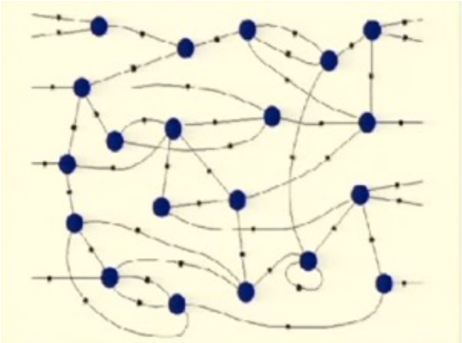
- Replicar a função biológica
- Replicar a estrutura biológica

Introdução

Inspiração biológica:

- Replicar a função biológica
- Replicar a estrutura biológica





Introdução

Função do algoritmo de aprendizado?

- Tem como função modificar os pesos sinápticos da rede, de maneira a ajusta-los para cumprir o objetivo desejado

- Modificação dos pesos sinápticos: método tradicional de aprendizado

- Também é possível que uma rede modifique sua própria topologia durante o aprendizado

- Motivação: neurônios no cérebro humano morrem, e novas conexões são criadas

Modelo de Neurônio

Neurônio: unidade de processamento de informação, fundamental para a operação da rede neural

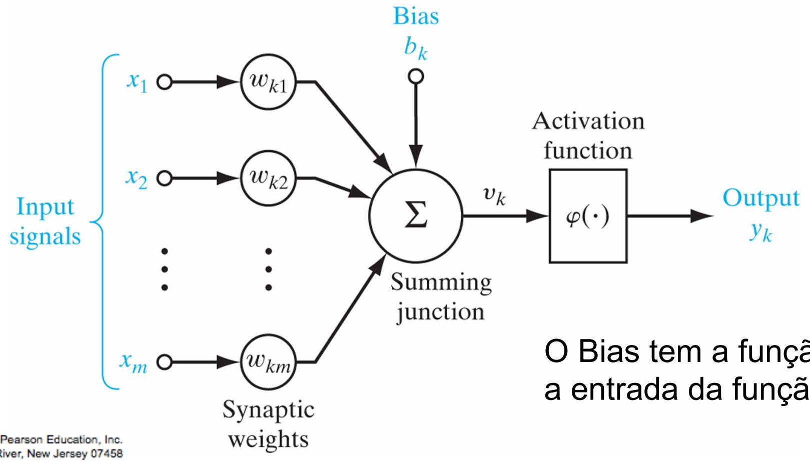
- Conjunto de sinapses, cada uma caracterizada por um peso: um sinal x_j na entrada da sinapse j conectada ao neurônio k, é multiplicado pelo peso sináptico w_{kj}. Os valores dos pesos podem ser negativos ou positivos

- Um somatório, que faz a combinação linear das entradas (soma as entradas ponderadas pelos seus pesos sinápticos)

- Função de ativação, para limitar a amplitude da saída do neurônio. Tipicamente [0,1] ou [-1,1]

Modelo de Neurônio

Neurônio: unidade de processamento de informação, fundamental para a operação da rede neural



O Bias tem a função de aumentar ou diminuir a entrada da função de ativação

Modelo de Neurônio

Matematicamente, descrevemos o neurônio k como:

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \quad \text{e} \quad y_k = \varphi(u_k + b_k)$$

em que:

$x_1, x_2, ..., x_m$ são sinais de entrada
$w_1, w_2, ..., w_3$ são pesos sinápticos do neurônio k
u_k é a combinação linear das entradas
b_k é o Bias
$\varphi(\cdot)$ é a função de ativação
y_k é a saída do neurônio

Modelo de Neurônio

- O uso do bias b_k tem o efeito de aplicar uma transformação em u_k, mantendo colinearidade

$$v_k = u_k + b_k$$

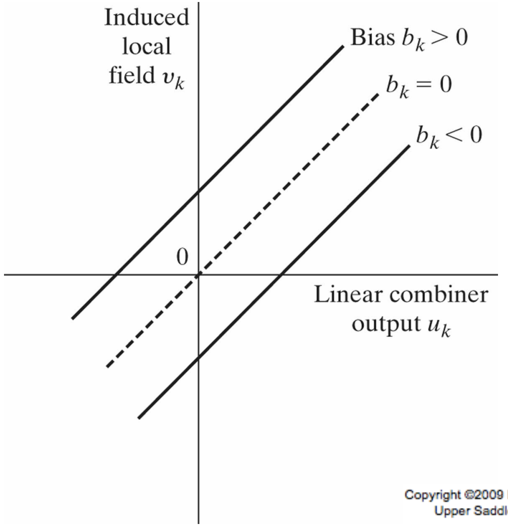
- Dependendo se b_k é positivo ou negativo, a relação entre o potencial de ativação v_k do neurônio k, e a saída da combinação linear u_k, é modificada

- O gráfico de v_k contra u_k não passa mais pela origem

Modelo de Neurônio

O uso do bias b_k tem o efeito de aplicar uma transformação em u_k, mantendo colinearidade

$$v_k = u_k + b_k$$



Modelo de Neurônio

- O bias b_k é um parâmetro externo do neurônio *k*. Ele pode ser incorporado à equação do neurônio

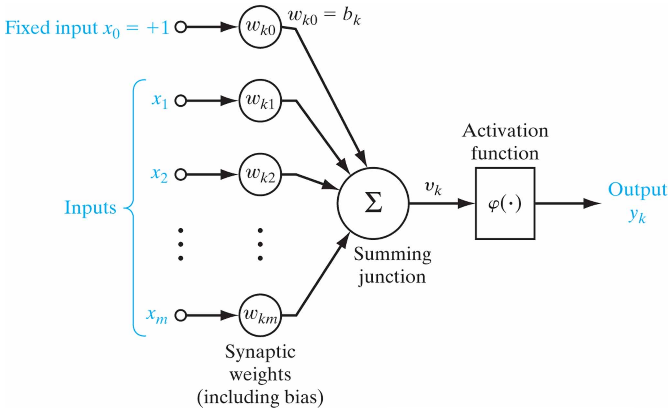
$$v_k = \sum_{j=0}^{m} w_{kj} x_j \qquad \text{e} \qquad y_k = \varphi(v_k)$$

- Simplesmente foi adicionada uma entrada x_0 com seu respectivo peso sináptico

$$x_0 = +1 \qquad \text{e} \qquad w_{k0} = b_k$$

Modelo de Neurônio

O bias b_k é um parâmetro externo do neurônio *k*. Ele pode ser incorporado à equação do neurônio. Assim chegamos a um modelo matematicamente equivalente



Funções de Ativação

- A função de ativação, denotada por $\varphi(v)$, define a saída do neurônio em termos de v:

Função degrau: $$\varphi(v) = \begin{cases} 1 & \text{se } v \geq 0 \\ 0 & \text{se } v < 0 \end{cases}$$

- A saída do neurônio *k* empregando essa função é dada por

$$y_k = \begin{cases} 1 & \text{se } v_k \geq 0 \\ 0 & \text{se } v_k < 0 \end{cases} \qquad \text{sendo} \qquad v_k = \sum_{j=1}^{m} w_{kj} x_j + b_k$$

Funções de Ativação

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

- Esse neurônio é conhecido como modelo de McCulloch-Pitts, devido ao trabalho de McCulloch e Pitts 1943

- A saída do neurônio é 1 se o potencial de ativação for positivo, e 0 caso contrário

Arquiteturas de Redes Neurais

A maneira como os neurônios são estruturados (arquitetura da rede) tem relação direta em como a rede neural é treinada

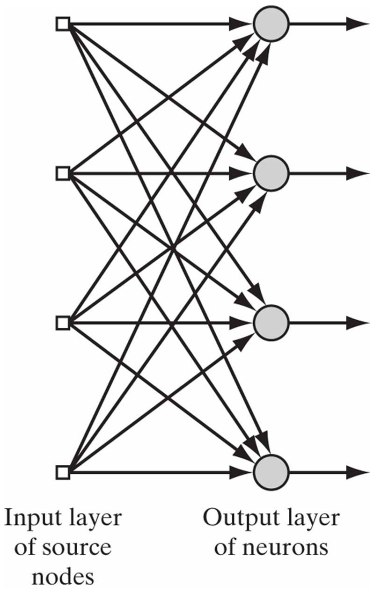
- Redes neurais de uma camada
- Redes neurais de múltiplas camadas
- Redes neurais recorrentes

Arquiteturas de Redes Neurais

Redes neurais de uma camada

- Neurônios organizados em camadas

- Camada de entrada de dados conectada a uma camada de neurônios de saída



Arquiteturas de Redes Neurais

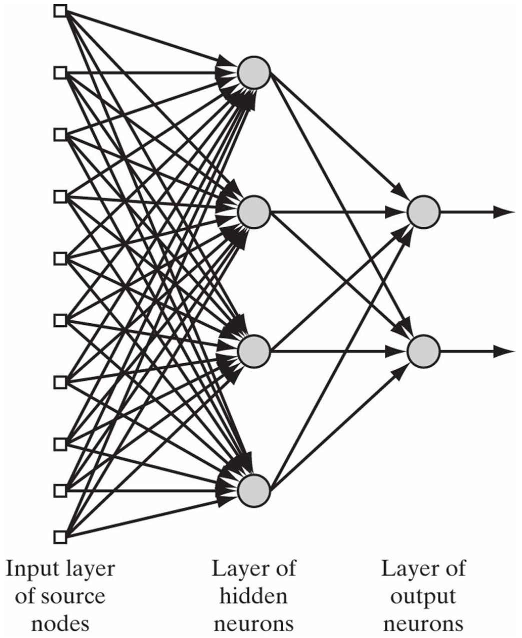
Redes neurais de múltiplas camadas

- Neurônios organizados em camadas
- Contém camadas escondidas atuam como extratores de estatísticas de mais alta ordem
- Neurônios de uma camada têm como entradas sinais provenientes apenas dos neurônios das camadas anteriores

Arquiteturas de Redes Neurais

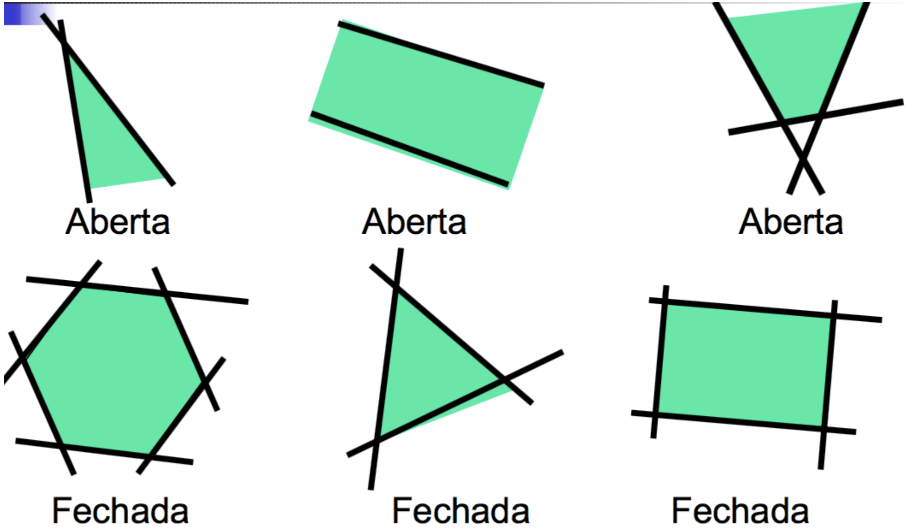
Redes neurais de múltiplas camadas

- Neurônios organizados em camadas
- Podem ser totalmente ou parcialmente conectadas



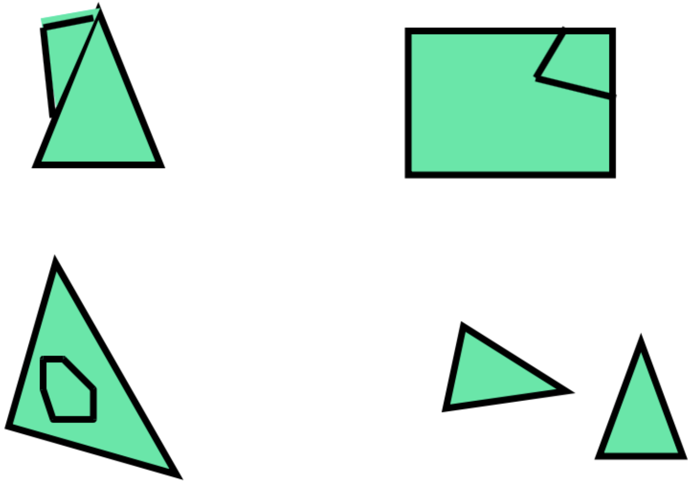
Arquiteturas de Redes Neurais

Múltiplas camadas - Regiões Convexas - Combinações de hiperplanos



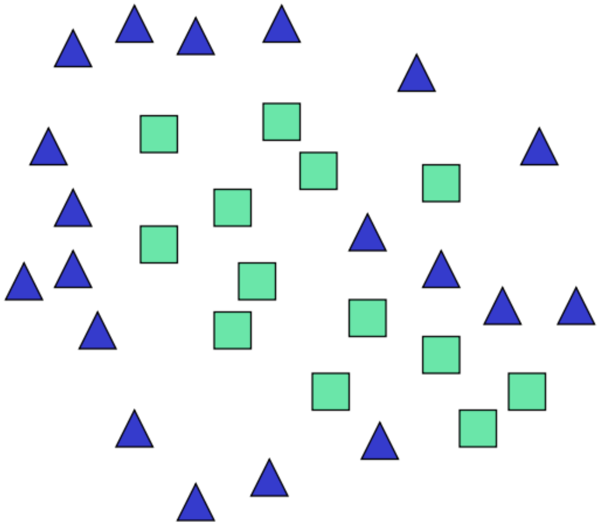
Arquiteturas de Redes Neurais

Múltiplas camadas – Combinações de regiões convexas



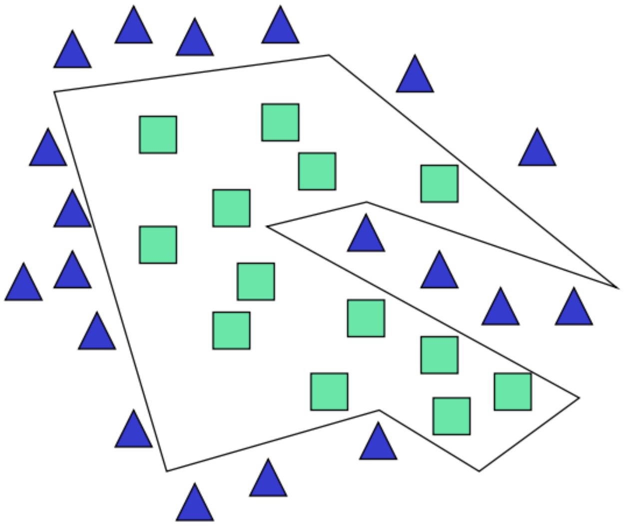
Arquiteturas de Redes Neurais

Múltiplas camadas - Regiões Convexas - Combinações de hiperplanos



Arquiteturas de Redes Neurais

Múltiplas camadas - Regiões Convexas - Combinações de hiperplanos



Representação do Conhecimento

- **Conhecimento**: informação armazenada, ou modelos usados por uma pessoa ou máquina para interpretar, predizer, e responder apropriadamente ao mundo externo

- Boas soluções dependem de boas representações do conhecimento

- A rede neural deve aprender um modelo do mundo (ambiente) no qual ele está, e manter esse modelo consistente com o mundo real para realizar uma tarefa (alcançar um objetivo)

Representação do Conhecimento

Conhecimento do mundo depende:

- 1. Conhecimento anterior: o estado do mundo, representado por fatos atuais e passados

- 2. Observações: obtidas por meio de sensores que sondam o ambiente no qual a rede neural vai operar

Observações podem conter ruído, devido a erros ou imperfeições na coleta. As observações coletadas são a fonte de informação, de onde se retirar os **exemplos** para treinar a rede neural.

Representação do Conhecimento

Como o conhecimento é produzido para a rede neural?

- ***Exemplos similares de classes similares devem produzir representações similares dentro da rede, e devem portanto serem classificados numa mesma classe***

- Como medir similaridade?

- Distância Euclidiana: $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{im}]^T$

Representação do Conhecimento

- O vetor $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{im}]^T$ define um ponto num espaço m-dimensional chamado espaço Euclidiano $\Re^m$

- A **distância Euclidiana** entre $\mathbf{x}_i$ e $\mathbf{x}_j$ é definida como:

$$d(\mathbf{x}_i, \mathbf{x}_j) = \|\mathbf{x}_i - \mathbf{x}_j\| = \left[\sum_{k=1}^{m} (x_{ik} - x_{jk})^2\right]^{1/2}$$

- Quanto menor a distância Euclidiana, maior é a similaridade entre os vetores, portanto mesma classe

Representação do Conhecimento

- Outra medida de similaridade: produto interno

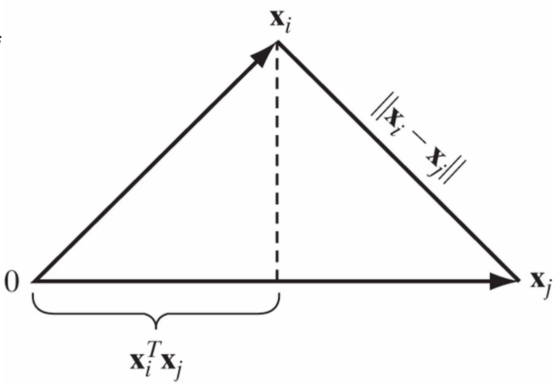
- O **produto interno** $\mathbf{x}_i^T \mathbf{x}_j$ entre dois vetores de mesma dimensão é definido como a projeção do vetor $\mathbf{x}_i$ no vetor $\mathbf{x}_j$

$$(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j = \sum_{k=1}^{m} x_{ik} x_{jk}$$

Representação do Conhecimento

- Essas duas medidas de similaridade são relacionadas. Quanto menor a distância Euclidiana, maior o produto interno

- Se normalizarmos os vetores, teremos $\|\mathbf{x}_i\| = \|\mathbf{x}_j$



$$d^2(\mathbf{x}_i, \mathbf{x}_j) = \|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2 - 2\mathbf{x}_i^T\mathbf{x}_j = 2 - 2\mathbf{x}_i^T\mathbf{x}_j$$

Perguntas?