

FIAP GRADUAÇÃO

ENGENHARIA DE COMPUTAÇÃO

Inteligência Artificial e Computacional

PROF. ANTONIO SELVATICI

SHORT BIO



É engenheiro eletrônico formado pelo Instituto Tecnológico de Aeronáutica (ITA), com mestrado e doutorado pela Escola Politécnica (USP), e passagem pela Georgia Institute of Technology em Atlanta (EUA). Desde 2002, atua na indústria em projetos nas áreas de robótica, visão computacional e internet das coisas, aliando teoria e prática no desenvolvimento de soluções baseadas em Machine Learning, processamento paralelo e modelos probabilísticos. Desenvolveu projetos para Avibrás, IPT, CESP e Systax.

PROF. ANTONIO SELVATICI

profantonio.selvatici@fiap.com.br

2. MACHINE LEARNING

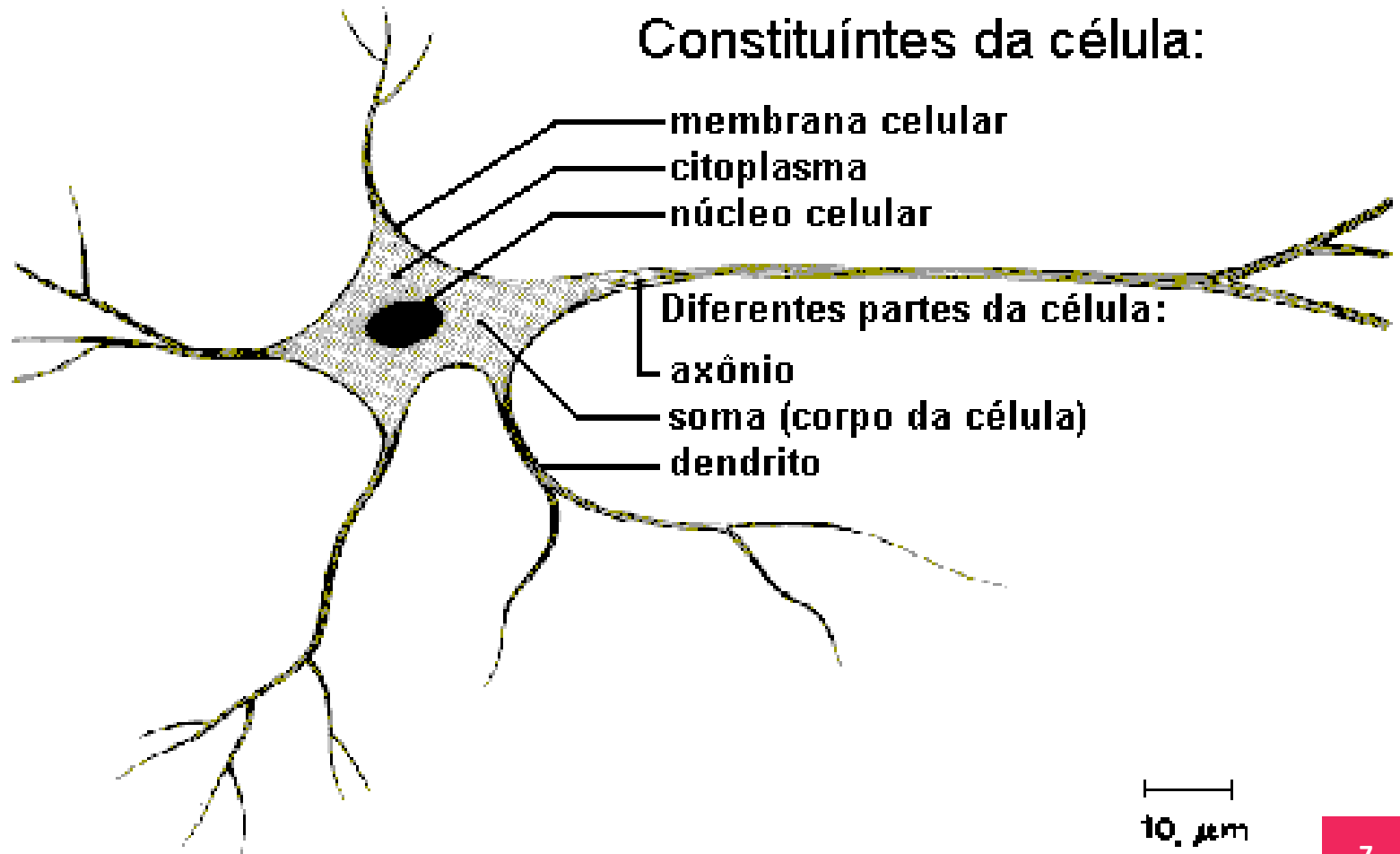
Redes Neurais Artificiais (RNA)

- O cérebro humano é considerado o mais fascinante processador baseado em carbono existente, sendo composto por aproximadamente 100 bilhões neurônios (86 bilhões segundo pesquisas recentes). Todas as funções e do organismo humano estão relacionados ao funcionamento destas pequenas células.
- As redes neurais são modelos computacionais inspirados pelo sistema nervoso de um animal capazes de realizar o aprendizado de máquina bem como o reconhecimento de padrões.
- Tais modelos são muitas vezes utilizados para a tarefa de classificação de padrões, podendo gerar classificadores com características variadas.
- As redes neurais artificiais possuem em comum o fato de serem constituídas por neurônios que se conectam entre si através de sinapses. No entanto, o significado particular do “neurônio” e da “sinapse” pode variar significativamente entre um modelo e outro.
- Exemplos de redes neurais artificiais usados em reconhecimento de padrões são a rede de Kohonem, as redes de base radial, a rede de Hopfield em principalmente, a rede Perceptron Multicamada (MLP)

O neurônio biológico

- O primeiro modelo de neurônio artificial foi o proposto pelo neurofisiologista McCulloch e pelo matemático Walter Pitts (1943), cujo trabalho fazia uma analogia entre células vivas e o processo eletrônico, simulando o comportamento do neurônio natural
- Os principais componentes dos neurônios são:
 - Os **dendritos**, que têm por função receber os estímulos transmitidos pelos outros neurônios;
 - O **corpo de neurônio**, também chamado de *soma*, que é responsável por coletar e combinar informações vindas de outros neurônios;
 - O **axônio**, que é constituído de uma fibra tubular que pode alcançar até alguns metros, e é responsável por transmitir os estímulos para outras células.

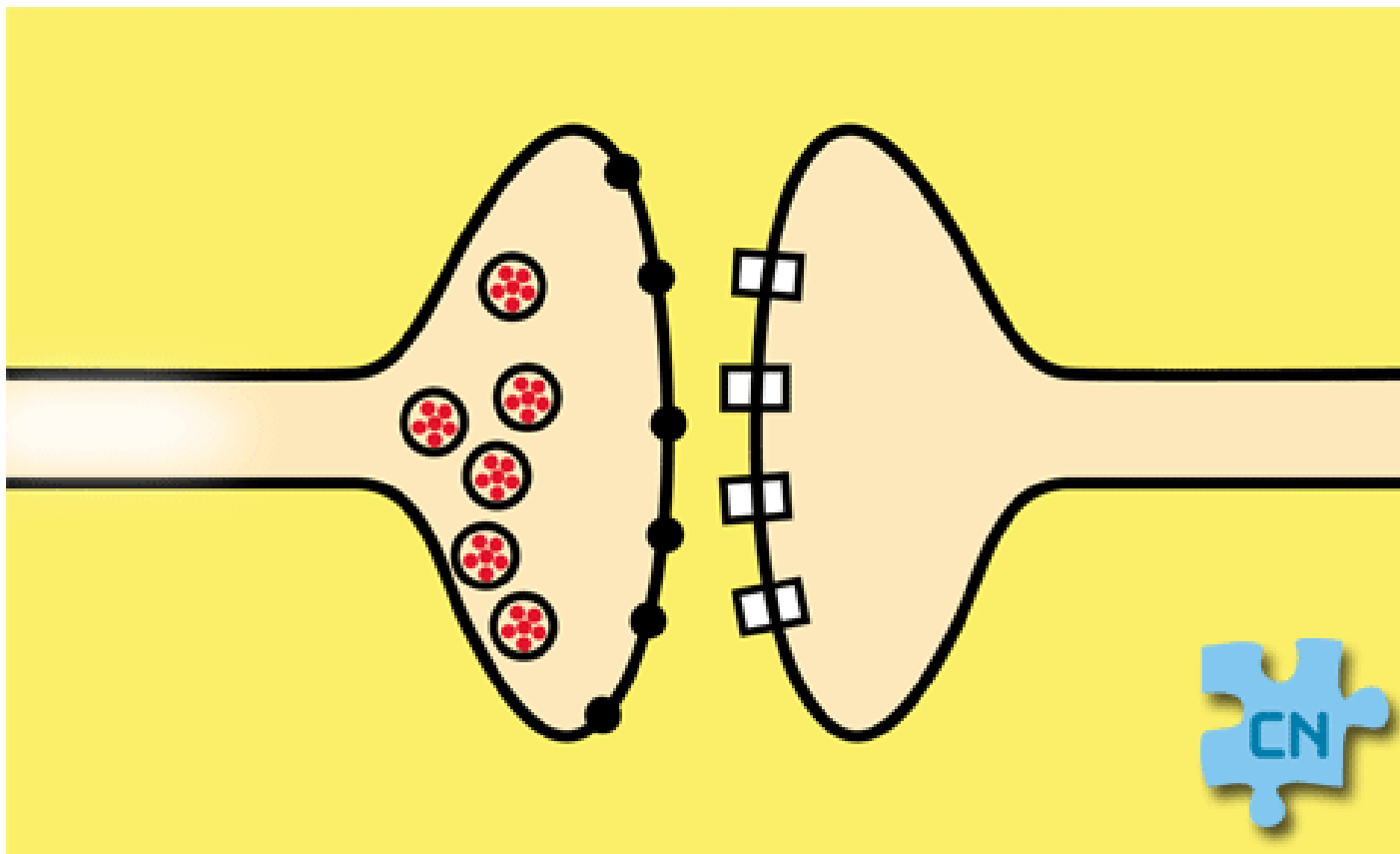
O neurônio biológico



Como funciona um neurônio natural?

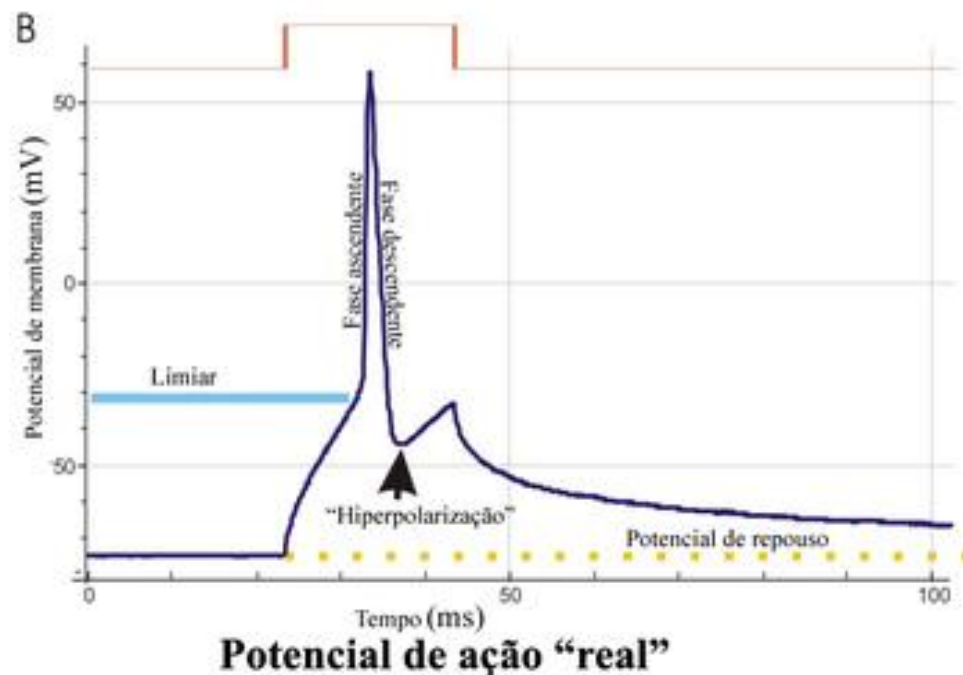
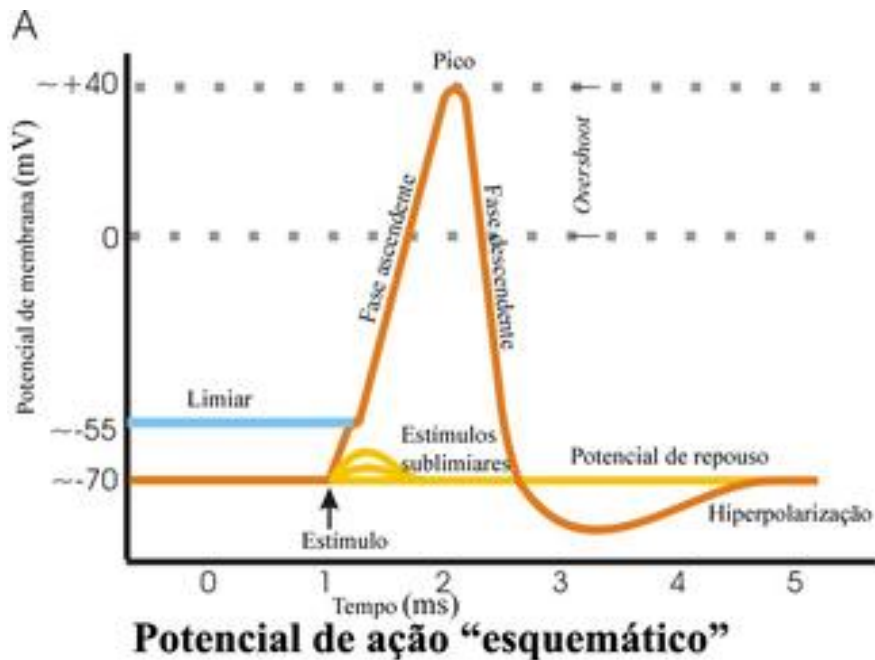
- Os neurônios são células conectadas umas às outras através de **sinapses**, que transmitem estímulos elétricos através de um processo bioquímico, envolvendo a emissão e a recepção de substâncias chamadas **neurotransmissores** ou **neuromoduladores**, dependendo da sua função
- As sinapses são resultado do encontro de um terminal do axônio de um neurônio com o terminal de um dendrito de outro neurônio
- O corpo do neurônio possui uma diferença de potencial com relação ao meio externo da célula; quando essa diferença de potencial atinge um valor denominado **potencial de ação** (algo em torno de 0.5V), as vesículas contendo os neurotransmissores ou neuromoduladores se fundem à membrana do axônio e liberam a substância que carregam. Essa liberação de substâncias pode ser interpretada como a transmissão de um pulso elétrico
- Dependendo do que foi liberado, os neurônio conectado ao primeiro acabam por aumentar ou diminuir seu potencial interno. No caso de acréscimo, um novo pulso pode ocorrer se o potencial de ação dessa célula for atingido

Transmissão sináptica



<http://www.cerebronosso.bio.br/sinapses/>

O comportamento do potencial do neurônio

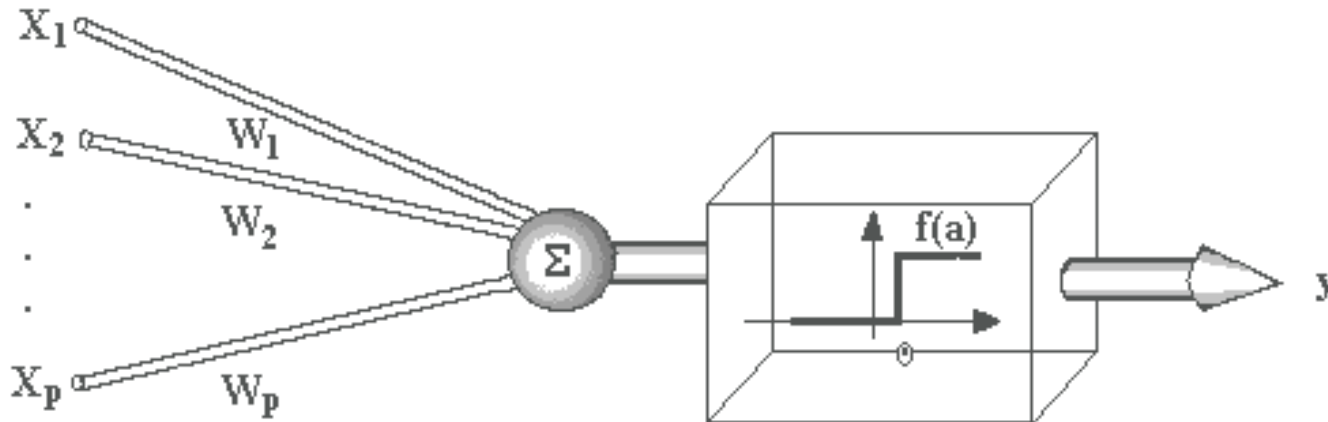


https://pt.wikipedia.org/wiki/Potencial_de_a%C3%A7%C3%A3o

O neurônio de McCulloch e Pitts

- O fisiologista Warren McCulloch interpretou o funcionamento do neurônio biológico como sendo um circuito elétrico com entradas binárias, ou seja, com valores em 0 ou 1, combinadas por uma soma ponderada produzindo uma entrada efetiva:
- A operação de uma unidade de processamento (neurônio), proposta por McCulloch e Pitts em 1943, pode ser resumida da seguinte maneira:
 - Sinais binários são apresentados à entrada;
 - Cada sinal é multiplicado por um número, ou peso, que indica a sua influência na saída da unidade;
 - É feita a soma ponderada dos sinais que produz um nível de atividade;
 - Se este nível de atividade exceder um certo limite (threshold) a unidade produz uma determinada resposta de saída.

Esquema do neurônio de McCulloch e Pitts



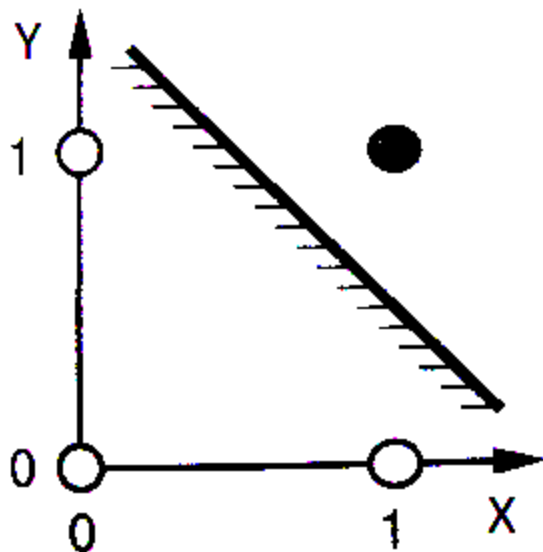
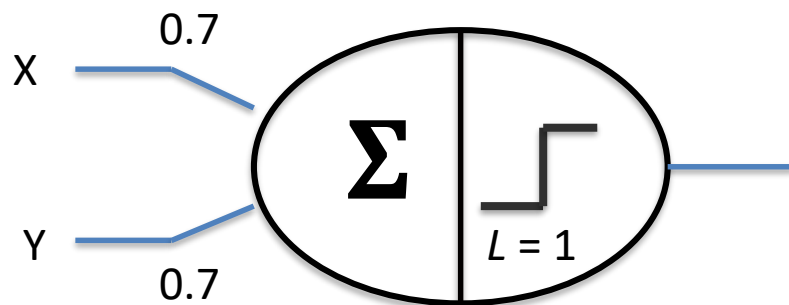
- x_1, x_2, \dots, x_p - Sinais de entrada
- w_1, w_2, \dots, w_p - Pesos que multiplicam cada entrada
- $a = w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_p \cdot x_p$ - Potencial interno do neurônio
- $y = f(a)$ - Saída do neurônio, com $f(.)$ sendo chamada de **função de ativação**

Na proposta original, $f(a)$ era uma função limiar, ou seja, de valor 0 (ou -1) para $a < L$, sendo L um **threshold** (limiar) predeterminado, ou 1 para $a \geq L$. Posteriormente, outras funções foram propostas para simular o comportamento de ativação do neurônio.

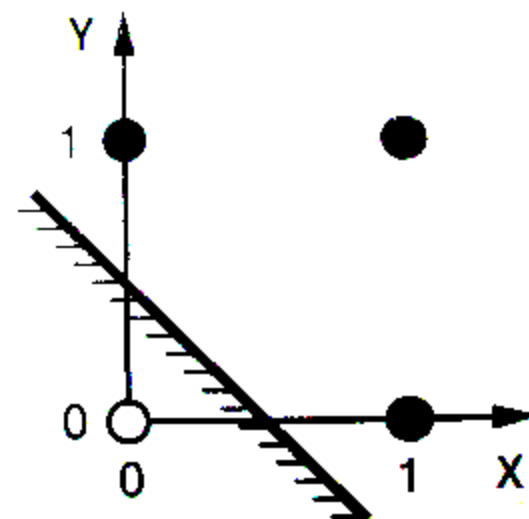
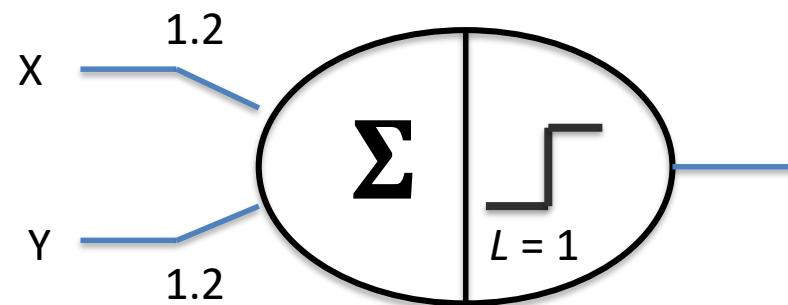
– Veja que não há o conceito de pulso elétrico, uma vez que, ocorrida a ativação, ela se torna um sinal constante (claro que se trata de uma simplificação)

Possíveis usos para o neurônio artificial

Porta AND (E)

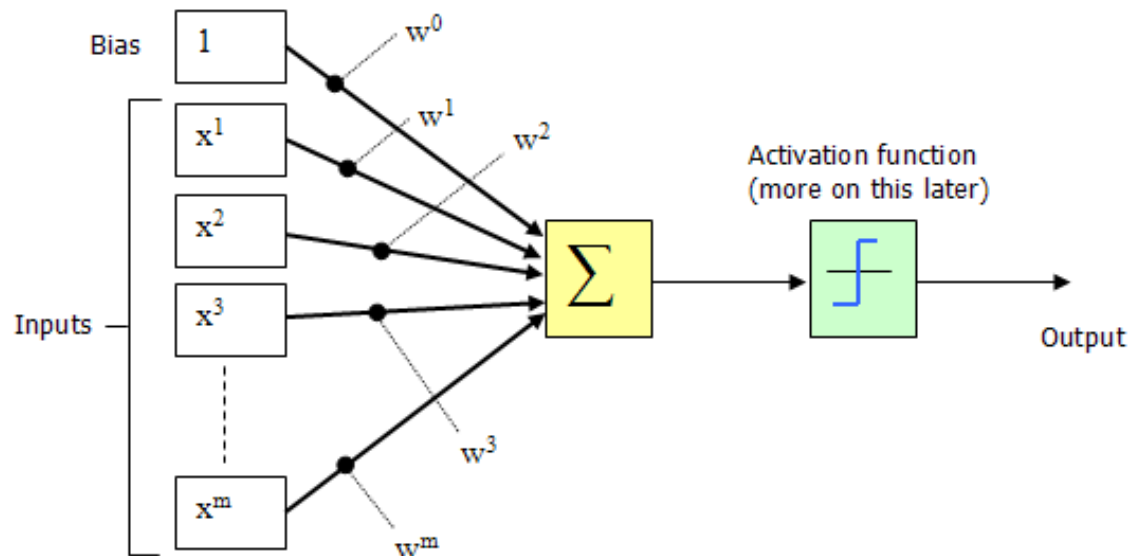


Porta OR (OU)



Bias do neurônio

- Uma forma de representarmos o limiar η separadamente, sem precisarmos trabalhar com a função de ativação, é criando um novo peso de bias (viés) na entrada do neurônio:
 - $w_{bias} = w_0 = -L$
 - O limiar da função de ativação passa a ser sempre zero

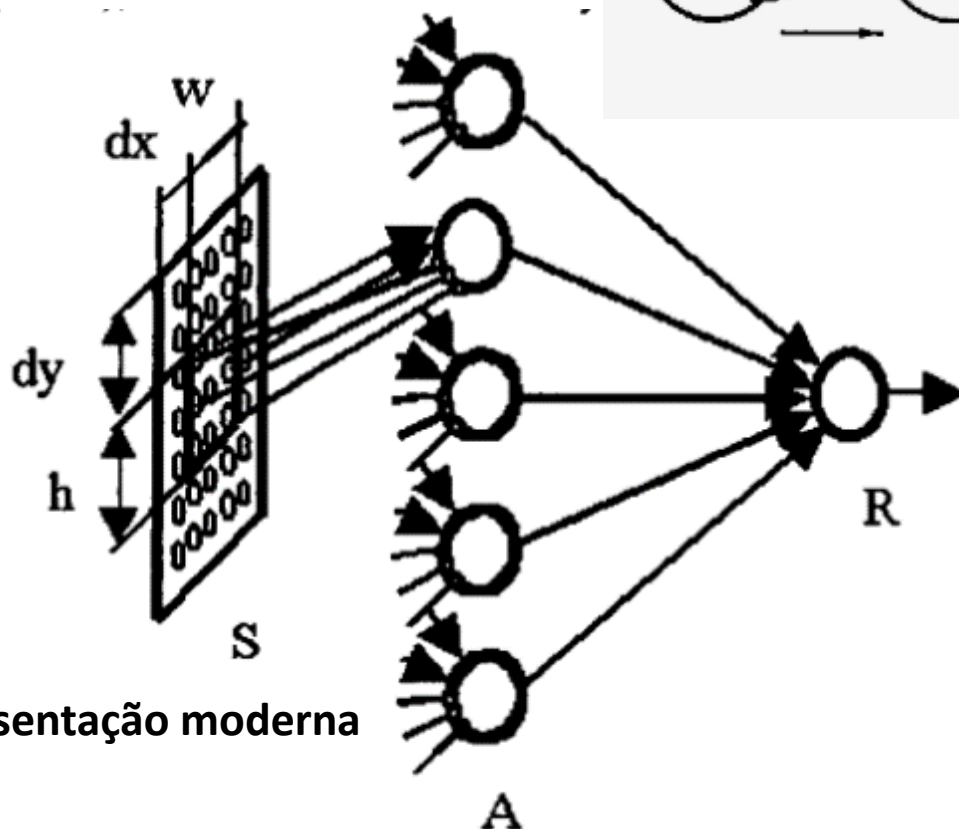
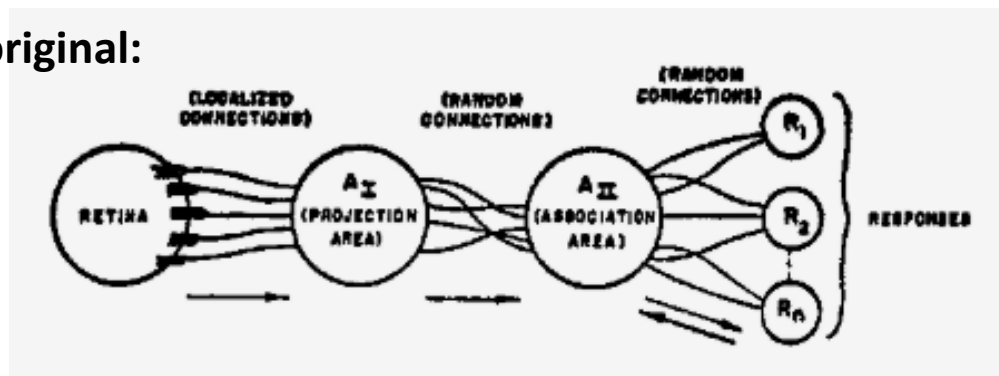


Redes de neurônios artificiais

- Em 1957, ROSENBLATT concebeu o Perceptron, que era uma rede neural de duas camadas* para processamento visual
 - *O número de camadas no artigo original é 4, mas torna-se apenas duas quando usamos uma notação moderna
- O Perceptron é atua como um classificador linear, ou seja, consegue separar duas classes em um espaço de atributos apenas se elas forem separáveis por uma reta, plano ou hiperplano
 - Dizemos que essas classes são linearmente separáveis
- A rede Perceptron possui um algoritmo de **aprendizado supervisionado** que consegue definir um classificador que encontra a superfície de separação entre quaisquer duas classes linearmente separáveis

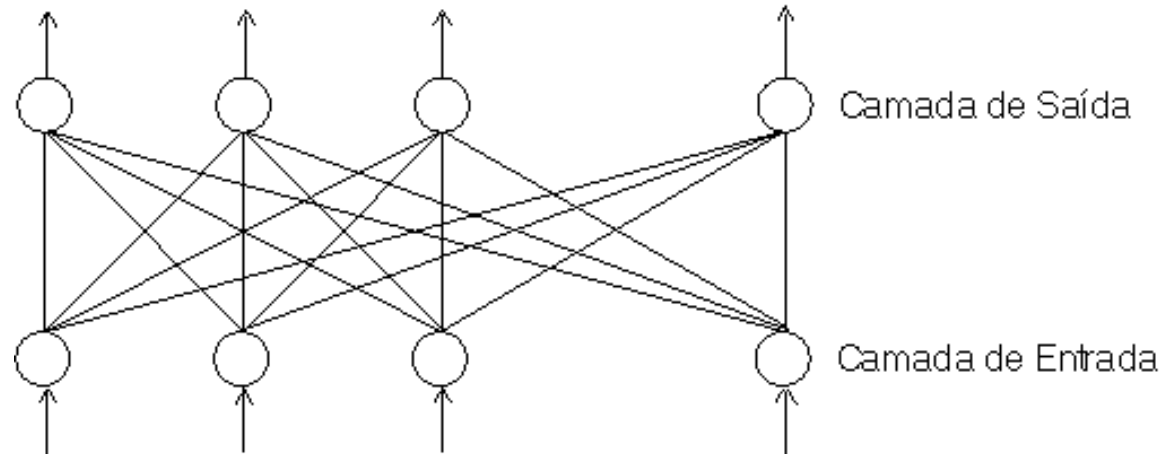
Representações da rede Perceptron

Artigo original:



Representação moderna

■ Perceptron

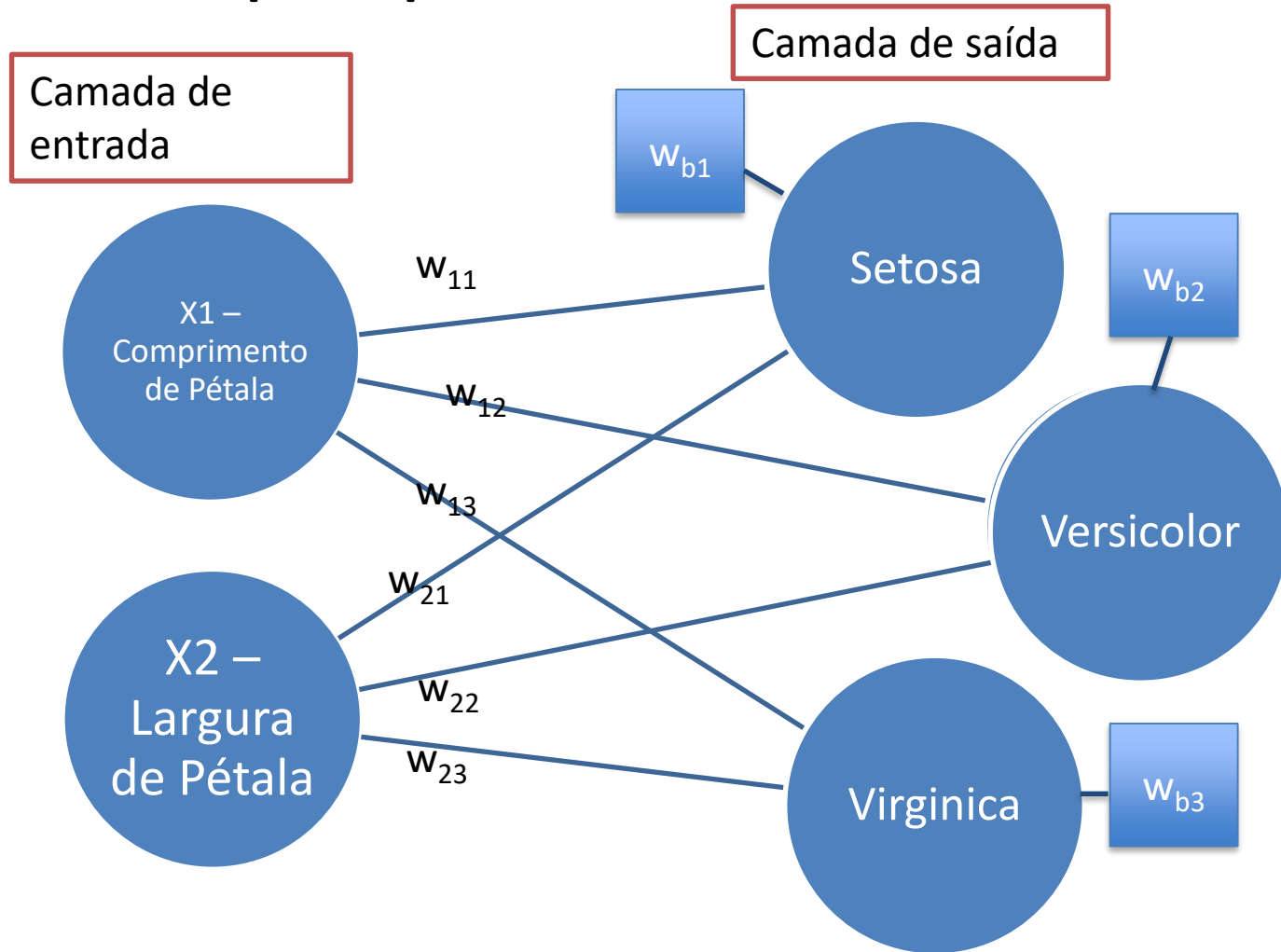


- A rede Perceptron é formada por duas camadas de neurônios :
 - Camada de entrada (S): onde os dados de entrada são recebidos
 - Cada atributo do vetor a ser classificado corresponde a um neurônio da entrada
 - Camada de saída (A/R): onde o resultado é gerado
- Como interpretar a camada de saída?
 - Cada neurônio de saída pode ser encarado como uma variável de um vetor
 - Dentro do contexto de classificação de padrões, caso houver múltiplas classes como opções para a classificação, cada neurônio representa uma classe. Assim, o neurônio com o maior valor de saída indica a classe escolhida no processo de classificação

Exercício

- Considere o conjunto de dados íris e a rede Perceptron que usa todos os atributos para classificar as espécies:
 - Quantos neurônios temos na camada de entrada?
 - Quantos neurônios na camada de saída?
 - Quantos pesos temos no total?
 - Desenhe a representação da rede

Rede Perceptron para o dataset Iris com 2 atributos



Regra de aprendizado do Perceptron

- Possuímos várias amostras rotuladas, classificadas por um especialista
- Para cada amostra rotulada, temos o vetor de atributos $v = (x_1, x_2, \dots, x_p)$ e classe c , sendo que a saída do neurônio j é dada por
 - $y_j = f(w_{bias,j} + w_{1,j} \cdot x_1 + w_{2,j} \cdot x_2 + \dots + w_{p,j} \cdot x_p)$
- Para cada amostra rotulada com a classe c , criamos o valor de alvo t_j do neurônio j , que vale 1 se a classe for igual à classe c_j , identificada pelo neurônio
 - $t_j = \begin{cases} 1, & \text{se } c = c_j \\ 0, & \text{se } c \neq c_j \end{cases}$
- Para cada amostra rotulada e para cada neurônio de saída j do Perceptron, atualizamos os pesos através da seguinte regra
 - $w_{i,j} \leftarrow w_{i,j} + \eta \cdot (t_j - y_j) \cdot x_i$, ou seja: $\Delta w_{i,j} = \eta \cdot e_j \cdot x_i$
 - η é chamada de “taxa de aprendizado”, deve ser um valor maior do que 0
- <https://www.youtube.com/watch?v=vGwemZhPlsA>

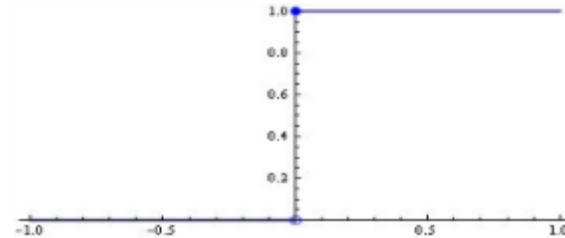
Medindo a intensidade da saída de um neurônio

- Como comparar a saída de diferentes neurônios ativados numa rede Perceptron?
 - Até agora, vimos neurônios com a função limiar, ou seja:
 - $f(a) = \begin{cases} 0, & a < 0 \\ 1, & a \geq 0 \end{cases}$
- E se trocássemos a função de ativação por outra que fosse contínua? Assim, além de determinar se um neurônio está ativou ou não, determinaríamos a “força” da ativação
 - Em neurônios biológicos, toda vez que um neurônio dispara, isso ocorre com a mesma intensidade.
 - Quando um neurônio está mais ativo, aumenta a frequência de pulsos, e não sua tensão ou corrente elétrica
 - Com isso, neurônios de saída mais ativos corresponderia a classes com maior probabilidade de casar com o dado sendo classificado

Diferentes funções de ativação

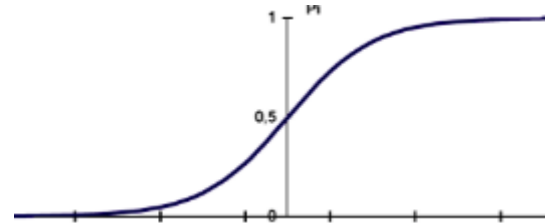
- Função limiar (ou Heaviside Step Function):

$$f(a) = \begin{cases} 0, & a < 0 \\ 1, & a \geq 0 \end{cases}$$



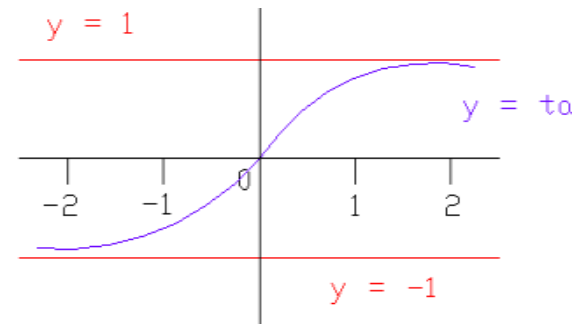
- Função logística

$$f(a) = \frac{1}{1+e^{-x}}$$



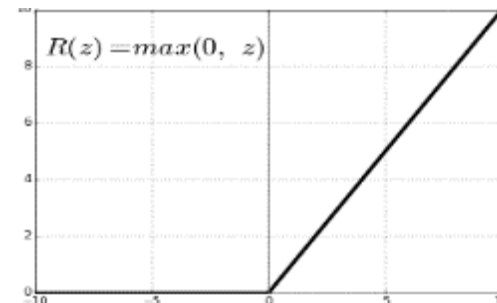
- Função tangente hiperbólica

$$f(a) = \tanh a$$



- Função Rectified Linear

$$f(a) = \begin{cases} 0, & a < 0 \\ a, & a \geq 0 \end{cases}$$

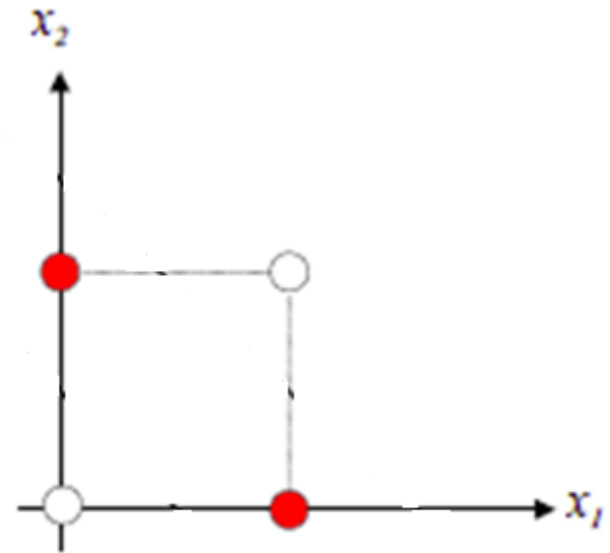


Problema da classificação XOR através do Perceptron

- Como encontrar uma reta que separe os pontos brancos dos pontos vermelhos?

- Tabela verdade XOR:

x_1	x_2	$x_1 \text{ XOR } x_2$
0	0	0
1	0	1
0	1	1
1	1	0



REFERÊNCIAS



- Stuart Russel & Peter Norvig. Inteligência Artificial – tradução da 2ª ed. Editora Campus, 2004
- Duda, Hart & Stork. Pattern Classification, 2nd. Ed., 2000
- Andre P. L. Carvalho. Redes Neurais Artificiais. <http://www.icmc.usp.br/~andre/research/neural/>
- Cassia Y. Tatibana & Deisi Y. Kaetsu. Redes Neurais. <http://www.din.uem.br/ia/neurais/>

Copyright © 2018 Prof. Antonio Selvatici

Todos direitos reservados. Reprodução ou divulgação total ou parcial deste documento é expressamente proibido sem o consentimento formal, por escrito, do Professor (autor).