



Inteligência Artificial

Profa. Patrícia R. Oliveira
EACH / USP

Parte 5 – Aprendizado Conexionista: Redes
Neurais Artificiais
(Modelos Perceptron)



Introdução

- Modelos que incorporam funções matemáticas (complexas).
- Podem ser usadas para a tarefa de classificação:
 - tomam uma instância como entrada e produzem uma saída, que é interpretada como a classe estimada pelo modelo.



Aprendizado de funções

- Aprendizado do mapeamento de instâncias em categorias (classes):
 - cada categoria é dada por um número.
 - ou por um intervalo de valores reais (por ex., 0.5 – 0.9).
- Exemplos de aprendizado de funções:
 - Entrada = 1, 2, 3, 4 Saída = 1, 4, 9, 16
 - Aqui, o conceito a ser aprendido é o quadrado dos números inteiros.
 - Entrada = [1,2,3], [2,3,4], [3,4,5], [4,5,6]
 - Saída = 1, 5, 11, 19
 - Aqui, o conceito é: $[a,b,c] \rightarrow a*c - b$

Exemplo: Classificando Veículos

- Entrada para a função: dados de pixels obtidos de imagens de veículos.
 - Saída: números: 1 para carro; 2 para ônibus; 3 para tanque

INPUT



OUTPUT = 3

INPUT



OUTPUT = 2

INPUT



OUTPUT = 1

INPUT



OUTPUT=1

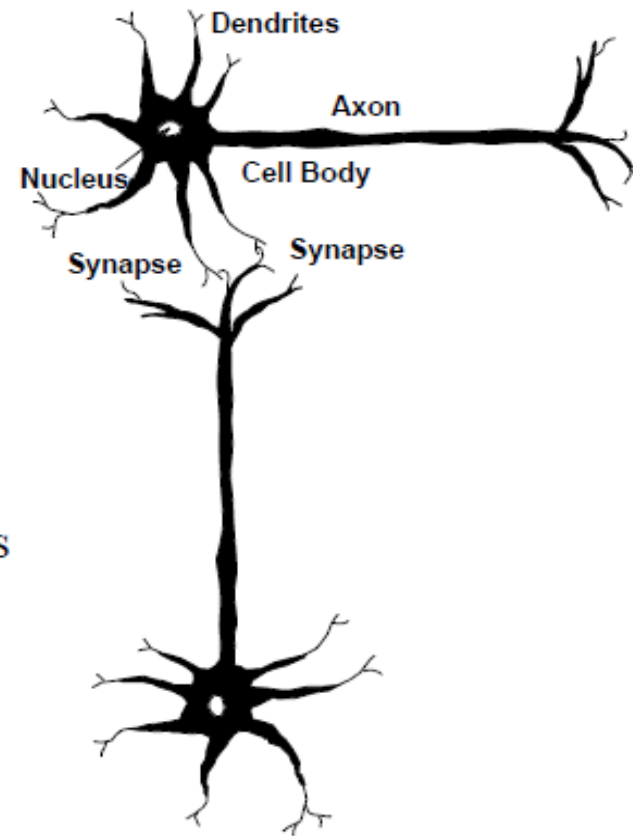


Por que usar Redes Neurais?

- **Motivação biológica:**
 - O cérebro faz com que tarefas de classificação pareçam fáceis.
 - O processamento cerebral é realizado por redes de neurônios.
 - Cada neurônio é conectado a vários outros neurônios.
- **Redes Neurais “Naturais”:**
 - A entrada de um neurônio é formada pelas saídas de vários outros neurônios.
 - Um neurônio é ativado se a soma ponderada de suas entradas $>$ limiar.

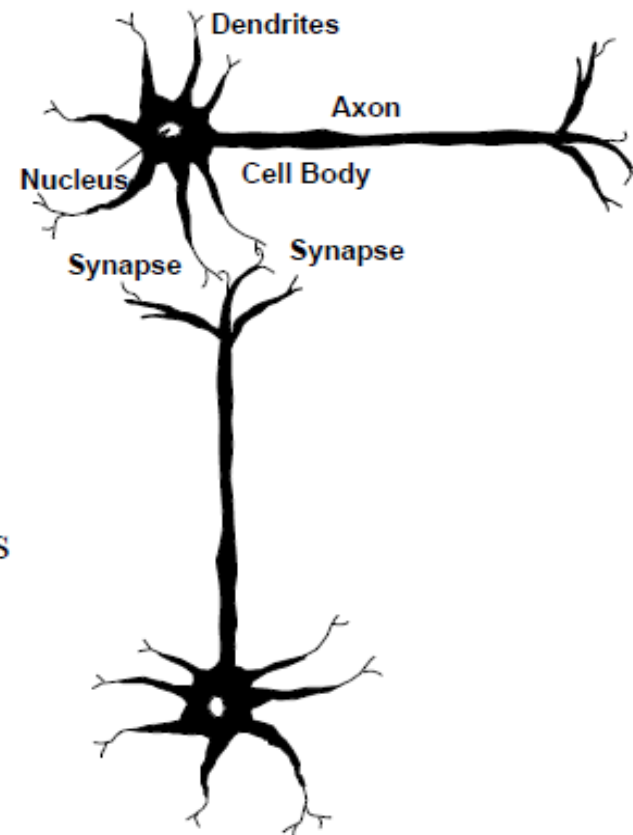
O neurônio biológico

- O neurônio recebe impulsos (sinais) de outros neurônios por meio dos seus dendritos.
- O neurônio envia impulsos para outros neurônios por meio do seu axônio.
- O axônio termina num tipo de contato chamado sinapse, que conecta-o com o dendrito de outro neurônio.



O neurônio biológico

- A sinapse libera substâncias químicas, chamadas de neurotransmissores, em função do pulso elétrico disparado pelo axônio.
- O neurônio envia impulsos para outros neurônios por meio do seu axônio.
- O fluxo de neurotransmissores nas sinapses pode ter um efeito excitatório ou inibitório sobre o neurônio receptor.





Processo de aprendizado

- O aprendizado ocorre por sucessivas modificações nas sinapses que interconectam os neurônios, em função da maior ou menor liberação de neurotransmissores.
- A medida que novos eventos ocorrem, determinadas ligações entre neurônios são reforçadas, enquanto outras são enfraquecidas.
- Este ajuste nas ligações entre os neurônios durante o processo de aprendizado é uma das mais importantes características das redes neurais artificiais.



Redes Neurais Artificiais (RNAs)

- Redes Neurais Artificiais (RNAs)
 - Hierarquia similar ao funcionamento do sistema biológico.
 - Neurônios que podem ser ativados por estímulos de entrada.
 - Mas essa analogia não vai muito longe...
 - Cérebro humano: aproximadamente 100.000.000.000 de neurônios.
 - RNAs: < 1000 geralmente

Ideia Geral

Valor calculado usando todos os valores das unidades de entrada

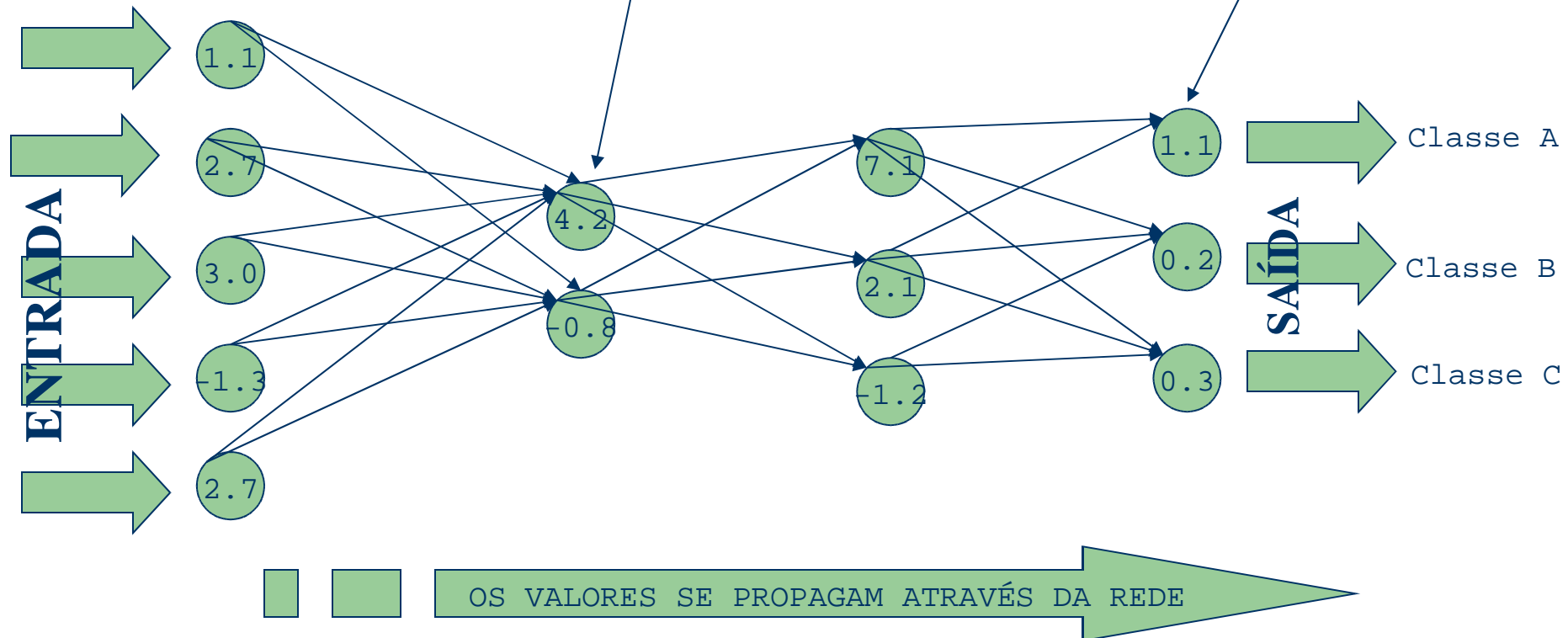
Escolhe a Classe A
(maior valor de saída)

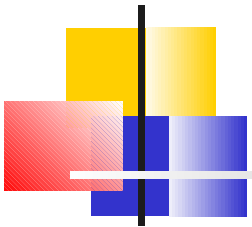
CAMADA DE ENTRADA

CAMADA OCULTA

CAMADA DE SAÍDA

CLASSE





Processamento das RNAs

- Cada unidade da rede realiza o mesmo cálculo.
 - geralmente baseado na soma ponderada das entradas na unidade.
- O conhecimento obtido pela rede fica armazenado nos pesos correspondentes a cada uma de suas unidades (neurônios).
- Representação “Caixa Preta”:
 - É difícil extrair o conhecimento sobre o conceito aprendido.



Aprendizado Supervisionado em RNAs

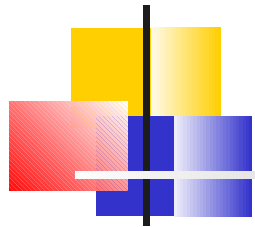
- Dados: conjunto de exemplos rotulados e representados numericamente.
- Tarefa: treinar uma rede neural usando esses exemplos.
 - O desempenho deve ser medido pela capacidade da rede em produzir saídas corretas para dados não contidos no conjunto de treinamento.



Aprendizado Supervisionado em RNAs

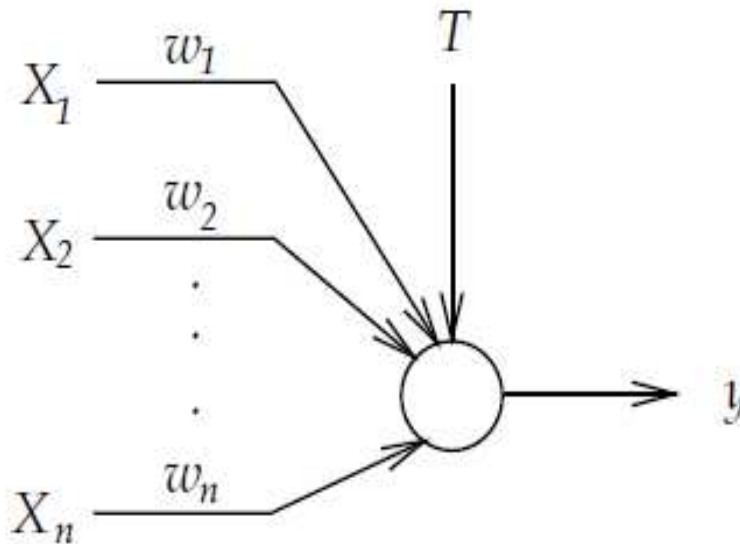
- Etapas preliminares ao treinamento:
 - escolha da arquitetura de rede correta.
 - número de neurônios
 - número de camadas ocultas
 - escolha da função de ativação (a mesma) para cada neurônio.
- A etapa de treinamento resume-se a:
 - ajustar os pesos das conexões entre as unidades para que a rede produza saídas corretas.

Perceptrons



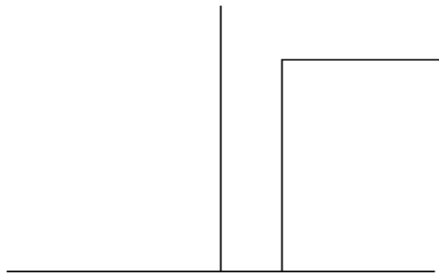
- O tipo mais simples de Rede Neural.
- Possui um único neurônio de saída.
 - Considera uma soma ponderada das entradas.
 - A função de ativação da unidade calcula a saída da rede.
 - Exemplo: unidade com threshold (limiar) linear.

$$\begin{aligned} - Netinput &= \sum_{i=1}^n x_i w_i \\ - if Netinput &\geq T \text{ then } y = 1 \text{ else } y = 0 \end{aligned}$$

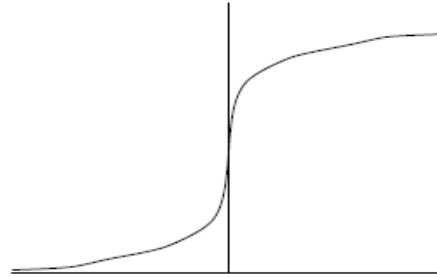


Perceptrons

- Algumas funções de transferência:



Função threshold



Função logística/sigmóide

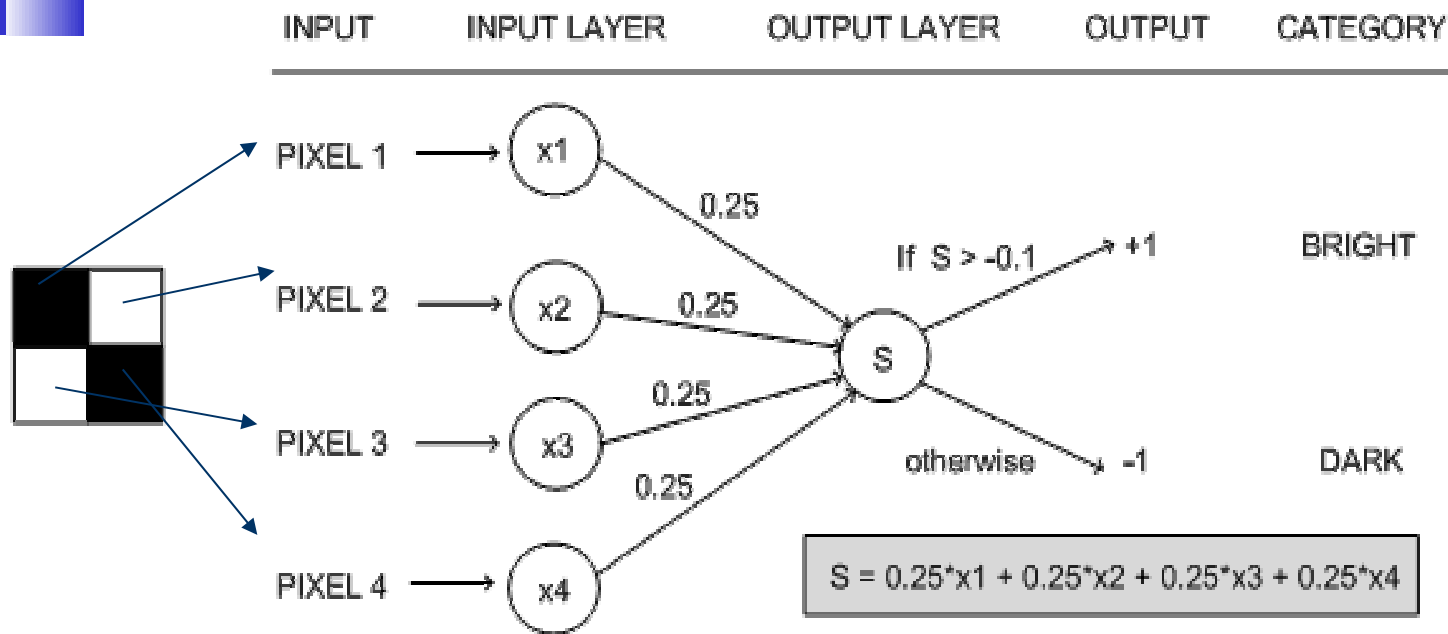
- Função Step (degrau):
 - Saída +1 se $Netinput > Threshold\ T$
 - Saída -1 caso contrário
 - Aqui, os dados binários são representados por +1 e -1
- Problema principal: como aprender os valores dos pesos da rede?



Perceptrons: Exemplo

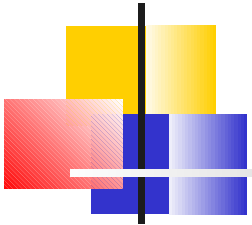
- Classificação de imagens preto e branco representadas por uma matriz de pixels 2x2.
 - Em “clara” ou “escura”.
- Pode-se representar o problema por essa regra:
 - Se apresentar 2, 3 ou 4 pixels brancos, é “clara”.
 - Se apresentar 0 ou 1 pixels brancos, é “escura”.
- Arquitetura do Perceptron:
 - Quatro unidades de entrada, uma para cada pixel.
 - Uma unidade de saída: +1 para “clara”, -1 para “escura”.

Perceptrons: Exemplo



- Exemplo de entrada: $x_1 = -1$, $x_2 = 1$, $x_3 = 1$, $x_4 = -1$
 - $S = 0.25 \cdot (-1) + 0.25 \cdot (1) + 0.25 \cdot (1) + 0.25 \cdot (-1) = 0$
- $0 > -0.1$, portanto a saída para a rede é +1
 - A imagem é classificada como "clara"

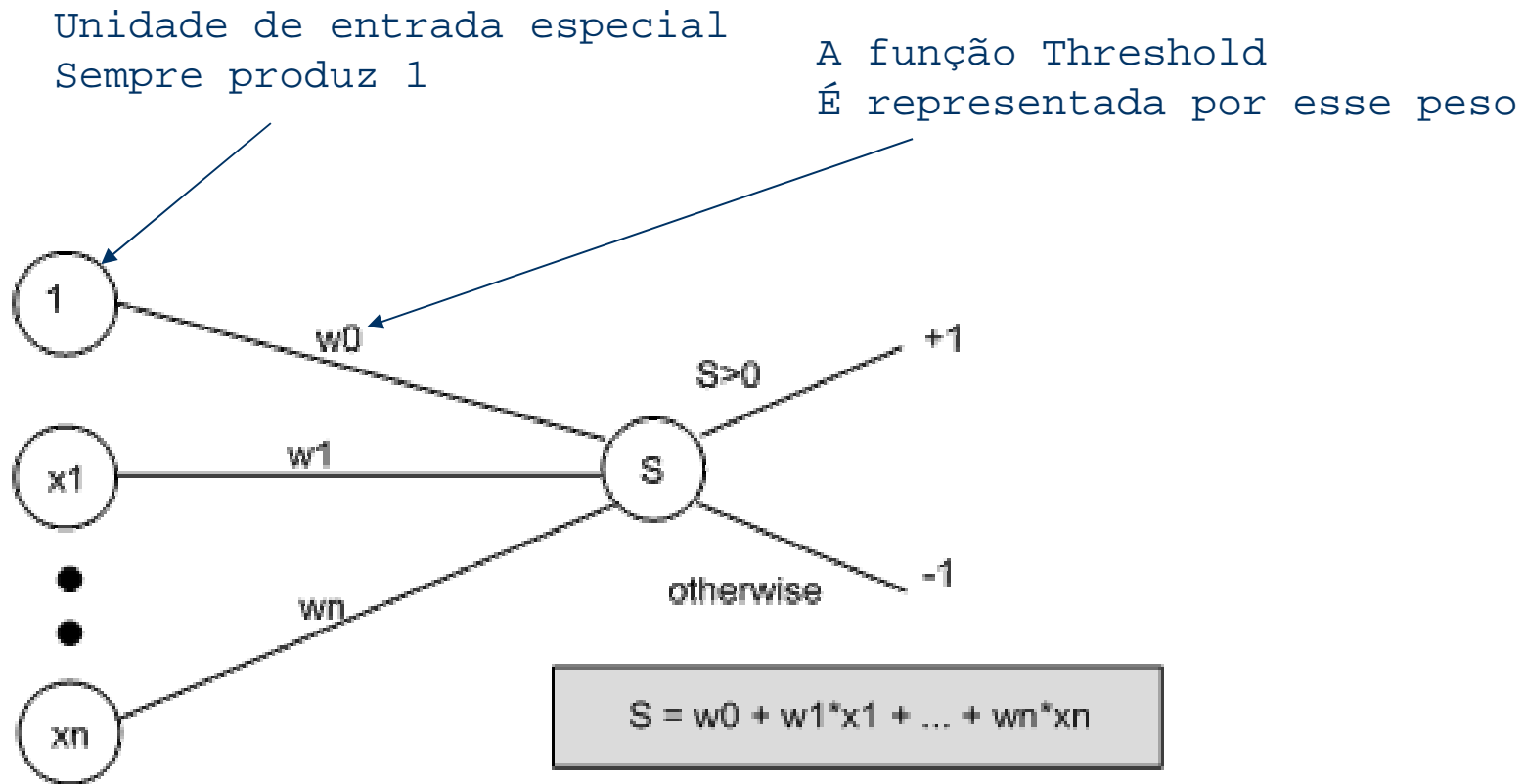
Aprendizagem em Perceptrons



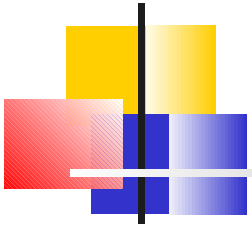
- É necessário aprender:
 - Os pesos entre as unidades de entrada e saída.
 - O valor do threshold.
- Para tornar os cálculos mais fáceis:
 - Considera-se o threshold como um peso referente a uma unidade de entrada especial, cujo sinal é sempre 1 (ou -1).
 - Agora, o único objetivo resume-se a aprender os pesos da rede.

Representação

Alternativa para Perceptrons

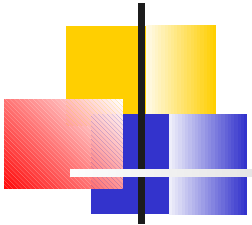


Perceptrons: Algoritmo de Aprendizagem (1)



- Os valores dos pesos são inicializados aleatoriamente, geralmente no intervalo $(-1, 1)$.
- Para cada exemplo de treinamento E :
 - Calcule a saída observada da rede $o(E)$.
 - Se a saída desejada $t(E)$ for diferente de $o(E)$:
 - Ajuste os pesos da rede, para que $o(E)$ chegue mais próximo de $t(E)$.
 - Isso é feito aplicando-se a regra de aprendizado do Perceptron.

Perceptrons: Algoritmo de Aprendizagem (2)



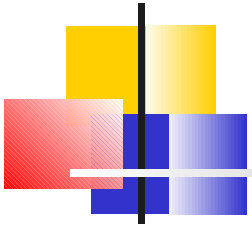
- O processo de aprendizado não para necessariamente depois de todos os exemplos terem sido apresentados.
 - Repita o ciclo novamente (uma “época”).
 - Até que a rede produza saídas corretas (ou boas o suficiente).
 - Considerando todos os exemplos no conjunto de treinamento.



Regra de Aprendizagem para Perceptrons

- Quando $t(E)$ for diferente de $o(E)$
 - Adicione Δ_i ao peso w_i
 - Em que $\Delta_i = \eta(t(E) - o(E))x_i$
 - Faça isso para todos os pesos da rede.

Regra de Aprendizagem para Perceptrons



- Interpretação:

$(t(E) - o(E))$ será igual a $+2$ ou -2

- Portanto, pode-se pensar na adição de Δ_i como uma movimentação do peso em uma determinada direção.
 - que irá melhorar o desempenho da rede com relação a E .
- Multiplicação por x_i
- O movimento aumenta proporcionalmente ao sinal de entrada.



Taxa de Aprendizado

- O parâmetro η é chamado de taxa de aprendizagem.
 - Geralmente escolhido como uma pequena constante entre 0 e 1 (por exemplo, 0.1).
- Controla o movimento dos pesos.
 - não deixa haver uma mudança grande para um único exemplo.
- Se uma mudança grande for mesmo necessária para que os pesos classifiquem corretamente um determinado exemplo:
 - Essa deve ocorrer gradualmente, em várias épocas.