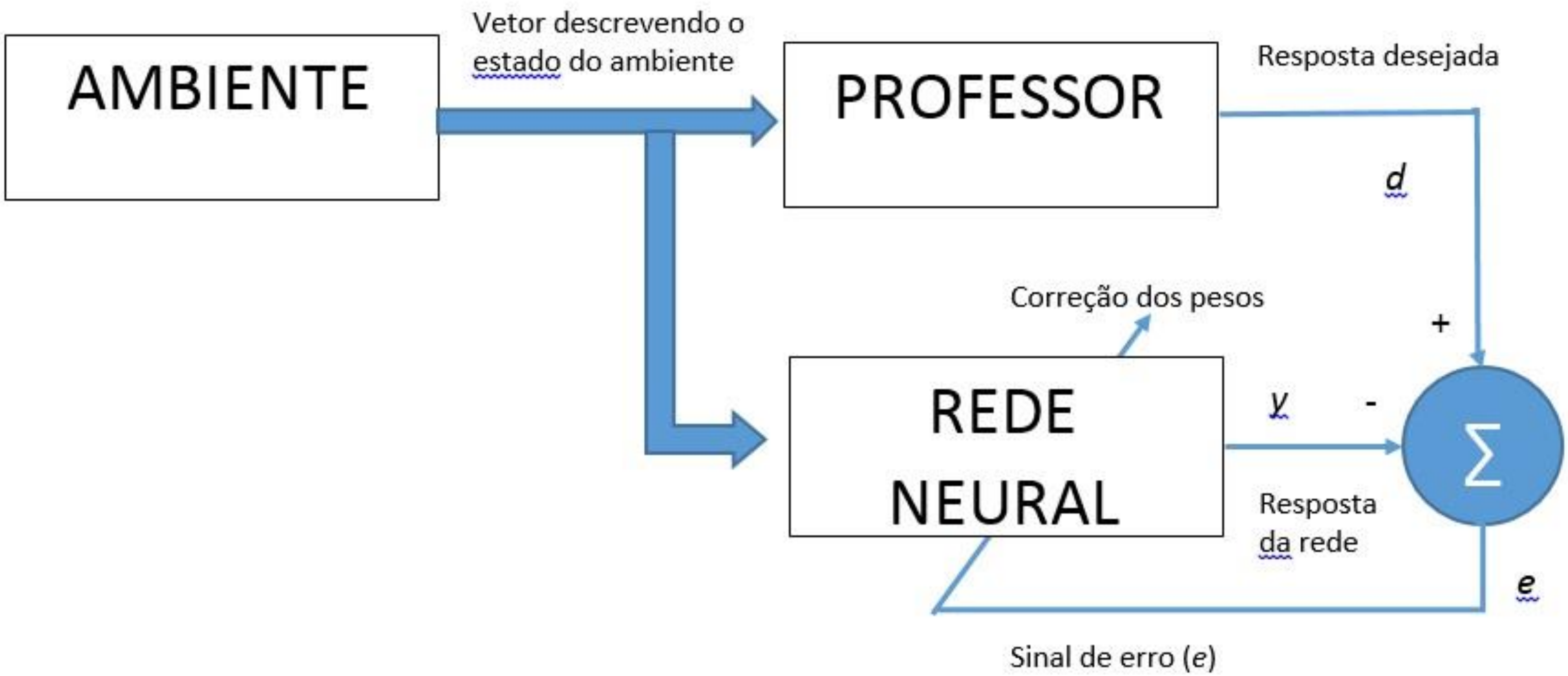
Aprendizado

- Determinação da intensidade das conexões (pesos) entre os neurônios através do algoritmo de aprendizagem.

Aprendizagem Supervisionada

- Conjunto de exemplos entrada-saída (professor)

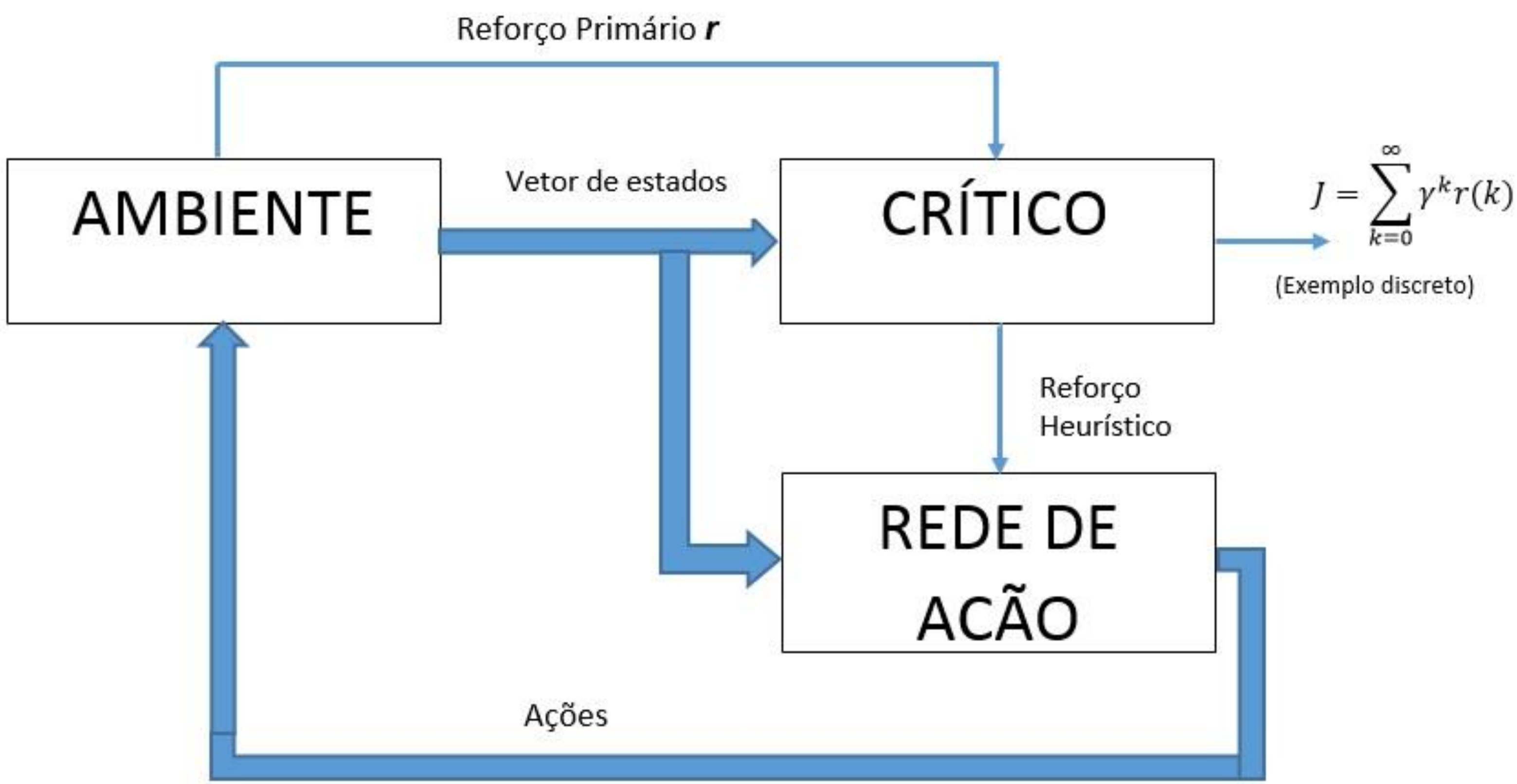


Aprendizagem não Supervisionada

- Não há professor (somente entradas disponíveis);
- A rede é sintonizada com as regularidades estatísticas dos dados de entrada, desenvolvendo a habilidade de formar representações internas para codificar caracteres de entrada e criar novas classes, ou grupos, que representam tais regularidades;
- Só é possível utilizar-se tal aprendizagem quando existe redundância nos dados de entrada (sem redundância seria impossível encontrar quaisquer padrões ou características dos dados de entrada);
- Exemplos: estágios iniciais da visão e audição.

Aprendizagem por Reforço (Sutton/Barto)

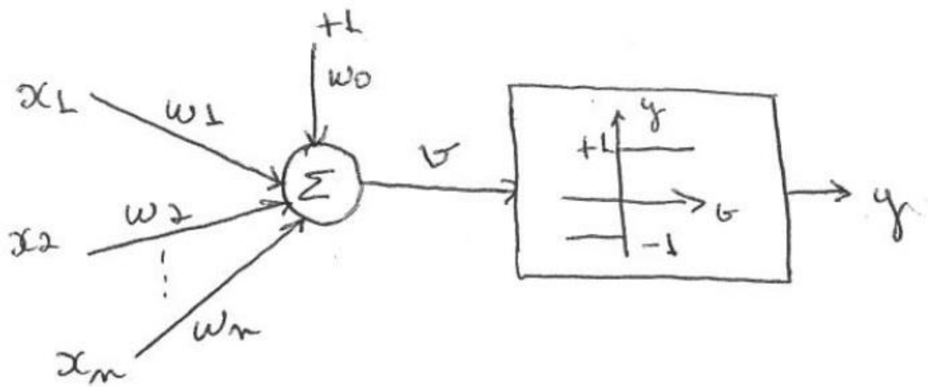
- Minimização/maximização de um critério de desempenho baseado em um sinal do ambiente (reforço primário: bom/ruim).



Aprendizagem supervisionada

- 1958→ Perceptron de Frank Rosemblatt;
- Teorema de convergência do perceptron mostra que um neurônio de McCulloch e Pitts treinado como o algoritmos de treinamento (aprendizagem) do perceptron sempre converge caso o problema seja linearmente separável.

Algoritmo de aprendizagem do perceptron



$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \Delta\boldsymbol{w}(n)\ \ (\boldsymbol{w} = vetor)$

$e = d - y\ \ , d = saída\ desejada\ para\ o\ padrão\ [x_1\ x_2\ \ ...\ x_n]^T$

$\Delta\boldsymbol{w} = \eta . e . \boldsymbol{X}$, (baseado na regra de Hebb originalmente)

$\eta = taxa\ de\ aprendizagem\ (influi\ no\ tempo\ de\ convergência)$

$0 \leq \eta \leq 1$, $\uparrow \eta$ -> *aprendizado oscila ou não converge*

$\downarrow \eta \ \rightarrow$ *demora mais para realizar o aprendizado*

Equação geral dos pesos do neurônio do perceptron:

$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \eta . e . \boldsymbol{X}\quad (\boldsymbol{w}, \boldsymbol{X} = vetor)$

Algoritmo

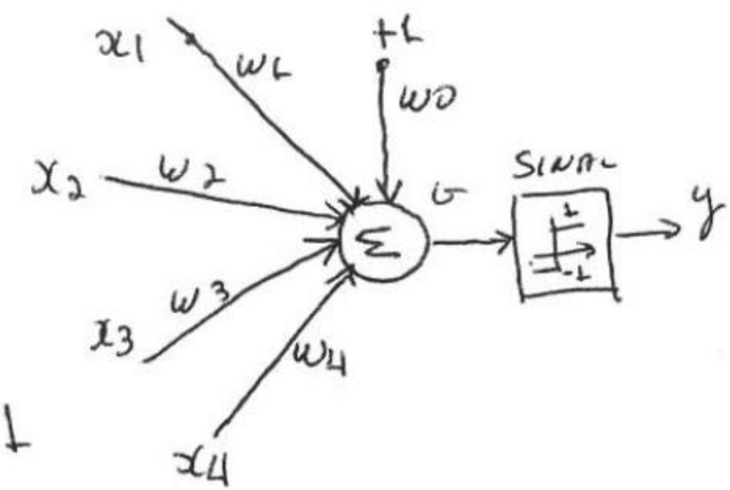
- Obs: 1 perceptron
  - 1- Inicializar η e w;
  - 2- Repetir
    - Para cada par do conjunto de treinamento $\Gamma = \{ X^i , y^i desejado\}_{i=1}^{p}$
    - Atualizar o vetor de pesos segundo a regra
      - $w(n+1) = w(n) + \eta e X(n)$; $X(n) = X^i$, $i=1$ a p *(atualiza padrão a padrão)*
      - ou
      - $w(n+1) = w(n) + \eta \sum_{i=1}^{p} e^i X^i$; *modo batelada- batch*
    - Até $e = 0$ para todos os p elementos do conjunto de treinamento

O algoritmo de treinamento do perceptron sempre chega a uma solução para o problema de separação de duas classes linearmente separáveis em um tempo finito (finitas iterações).

Exemplo

Aplicação a um problema de reconhecimento de padrões (2 classes linearmente separáveis)

Exemplo com 4 entradas treinado para separar 2 classes representadas pelos vetores



$$\left\{\begin{array}{l} \begin{array}{cccccc} & & x_1 & x_2 & x_3 & x_4 \end{array} \\ \mathbf{x}^1 = [+1\ \ -0{,}1\ \ +0{,}4\ \ -0{,}7\ \ -1{,}8]^T,\ d^1 = 1 \\ \mathbf{x}^2 = [+1\ \ 0{,}1\ \ 0{,}2\ \ -0{,}3\ \ -0{,}9]^T\ \ d^2 = -1 \end{array}\right.$$

apenas 2 vetores → linearmente separáveis

Exemplo (cont.)

Supondo $\eta = 0{,}03$ e $W(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ e MODO BATELADA

Exemplo (n=0)

- Para $n=0$: $W(1) = W(0) + \eta \sum_{i=1}^{2} e^i x^i$

$$i=1 \Rightarrow x^1 \Rightarrow e^1 = d_1 - y = 1 - y(0+\cancel{0}+0+0+0) = 1-1=0$$

1

$$i=2 \Rightarrow x^2 \Rightarrow e^2 = d_2 - y = -1 - y(\cancel{0}) = -1-1=-2$$

$+1$

$$W(1) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 0{,}03 \times \left\{ 0 \times \begin{bmatrix} 1 \\ -0{,}1 \\ 0{,}4 \\ -0{,}7 \\ -1{,}8 \end{bmatrix} - 2 \times \begin{bmatrix} 1 \\ 0{,}1 \\ 0{,}2 \\ -0{,}3 \\ -0{,}9 \end{bmatrix} \right\} = \begin{bmatrix} -0{,}06 \\ -0{,}006 \\ -0{,}012 \\ +0{,}018 \\ +0{,}054 \end{bmatrix}$$

$e^1 = 0$, $e^2 \neq 0 \Rightarrow$ continua!

Exemplo (n=1)

- Para n = 1 : $w(2) = w(1) + \eta \sum_{i=1}^{2} e^i x^i$

$$i=1 \Rightarrow x^1 \Rightarrow e^1 = 1 - y\left[-0{,}06(1) - 0{,}006(-0{,}1) - 0{,}012\cancelto{-1}{(0{,}4)} + 0{,}018(-0{,}7) + 0{,}054(-1{,}8)\right]$$

$$e^1 = 1 - (-1) = 2 \quad \text{(CONTINUA!)}$$

$$i=2 \Rightarrow x^2 \Rightarrow e^2 = -1 - y\left[-0{,}06(1) - 0{,}006(0{,}1) - 0{,}012\cancelto{-1}{(0{,}2)} + 0{,}018(-0{,}3) + 0{,}054(-0{,}9)\right]$$

$$e^2 = -1 - (-1) = 0$$

$$W(2) = \begin{bmatrix} -0{,}06 \\ -0{,}006 \\ -0{,}012 \\ 0{,}018 \\ 0{,}054 \end{bmatrix} + 0{,}03 \left\{ 2 \times \begin{bmatrix} 1 \\ -0{,}1 \\ 0{,}4 \\ -0{,}7 \\ -1{,}8 \end{bmatrix} + 0 \right\} = \begin{bmatrix} 0 \\ -0{,}12 \\ 0{,}012 \\ -0{,}024 \\ -0{,}054 \end{bmatrix}$$

Exemplo (n=2)

- Caso $n=2$: $w(3) = w(2) + \eta \sum_{i=1}^{2} e^i x^i$

$i=1 \Rightarrow x^1 \to y^1 = 1 - y\,[0(1) - 0{,}12(-0{,}1) + 0{,}012(0{,}4) - 0{,}024(-0{,}7) - 0{,}054(-1{,}8)] \to 1$

$e^1 = 1 - 1 = 0$

$i=2 \Rightarrow x^2 \Rightarrow e^2 = -1 - [0(1) - 0{,}12(0{,}1) + 0{,}012(0{,}2) - 0{,}024(-0{,}3) - 0{,}054(-0{,}9)] \to 1$

$e^2 = -1 - 1 = -2$ (CONTINUA!)

$$W(3) = \begin{bmatrix} 0 \\ -0{,}12 \\ 0{,}012 \\ -0{,}024 \\ -0{,}054 \end{bmatrix} + 0{,}03 \times \left\{ 0 - 2 \begin{bmatrix} 1 \\ 0{,}1 \\ 0{,}2 \\ -0{,}3 \\ -0{,}9 \end{bmatrix} \right\} = \begin{bmatrix} -0{,}06 \\ -0{,}126 \\ 0 \\ -0{,}006 \\ 0 \end{bmatrix}$$

Exemplo (n =3)

- para $n=3$: $W(4) = W(3) + \eta \sum_{i=1}^{2} e^i x^i$

$i=1 \Rightarrow x^1 \Rightarrow e^1 = 1 - y\,[-0{,}06\,(1) - 0{,}126\,(-0{,}1) + 0 - 0{,}006\,(-0{,}7) + 0] \to -1$

$e^1 = 1-(-1) = 2$ (CONTINUA!)

$i=2 \Rightarrow x^2 \Rightarrow e^2 = -1 - y\,[-0{,}06\,(1) - 0{,}126\,(0{,}1) + 0 - 0{,}006\,(-0{,}3) + 0] \to -1$

$e^2 = -1 - (-1) = \emptyset$

$$W(4) = \begin{bmatrix} -0{,}06 \\ -0{,}126 \\ 0 \\ -0{,}006 \\ 0 \end{bmatrix} + 0{,}03 \left\{ 2 \begin{bmatrix} 1 \\ -0{,}1 \\ 0{,}4 \\ -0{,}7 \\ -1{,}8 \end{bmatrix} + 0 \right\} = \begin{bmatrix} 0 \\ -0{,}132 \\ 0{,}024 \\ -0{,}048 \\ -0{,}108 \end{bmatrix}$$

Exemplo (n=4)

- para $n=4$: $w(5) = w(4) + \eta \sum_{i=1}^{2} e^i x^i$

$i = 1 \Rightarrow x^1 \Rightarrow e^1 = 1 - y[0(1) - 0{,}132(-0{,}1) + 0{,}024(0{,}4) - 0{,}048(-0{,}7) - 0{,}108(-1{,}8)]$

$e^1 = 1 - 1 = 0$

$i = 2 \Rightarrow x^2 \Rightarrow e^2 = -1 - y[0(1) - 0{,}132(0{,}1) + 0{,}024(0{,}2) - 0{,}048(-0{,}3) - 0{,}108(-0{,}9)]$

$e^2 = -1 - 1 = -2$ (CONTINUA)

$$w(5) = \begin{bmatrix} 0 \\ -0{,}132 \\ 0{,}024 \\ -0{,}048 \\ -0{,}108 \end{bmatrix} + 0{,}03 \left\{ 0 - 2 \begin{bmatrix} 1 \\ 0{,}1 \\ 0{,}2 \\ -0{,}3 \\ -0{,}9 \end{bmatrix} \right\} = \begin{bmatrix} -0{,}06 \\ -0{,}138 \\ 0{,}012 \\ -0{,}030 \\ -0{,}054 \end{bmatrix}$$

Exemplo (fim)

- para $n=5$: $W(6) = W(5) + \eta \sum_{l=1}^{2} e^l x^l$

$$l=1 \Rightarrow x^1 \Rightarrow e^1 = 1 - y\overbrace{[-0{,}06(1) - 0{,}138(-0{,}1) + 0{,}012(0{,}4) - 0{,}030(-0{,}7) - 0{,}054(-1{,}8)]}^{+1}$$

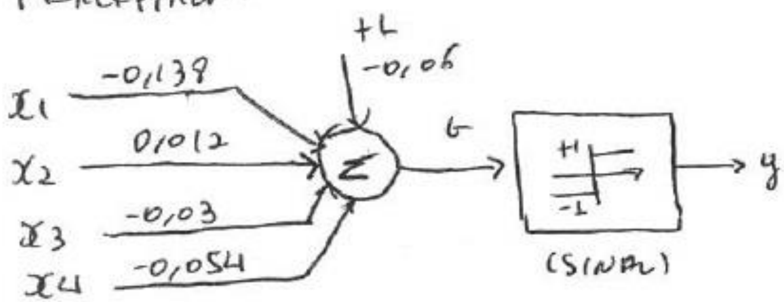
$$e^1 = 1 - 1 = 0$$

$$l=2 \Rightarrow x^2 \Rightarrow e^2 = -1 - y\overbrace{[-0{,}06(1) - 0{,}138(0{,}1) + 0{,}012(0{,}2) - 0{,}030(-0{,}3) - 0{,}054(-0{,}9)]}^{-1}$$

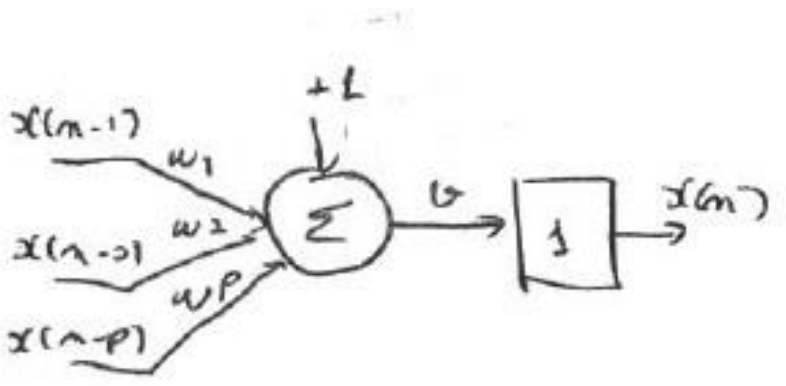
$$e^2 = -1 + 1 = 0$$

$e^1, e^2 = 0 \Rightarrow$ FIM $\boxed{\text{TREINOU COM 5 ITERAÇÕES}}$

PERCEPTRON:



Algoritmo LMS



- Neurônio Linear;
- Mesma regra de aprendizagem do perceptron;
- Aplicações: predição, filtragem adaptativa;
- Baseado em algoritmo de correção de erro;
- Aprendizagem contínua.

Redes Perceptron Multicamada (MLP)

- MLP→ "Multi-Layer Perceptron;
- Rede tipo *Feedforward*;
- Padrões não precisam ser linearmente separáveis;
- Podemos ter *n* camadas escondidas;
- Como treinar os pesos das camadas intermediárias? R: algoritmo de retropropagação do erro (*backpropagation*)→ próxima aula!