

REDES NEURAIS

A arquitetura da rede Multilayer Perceptron (MLP)

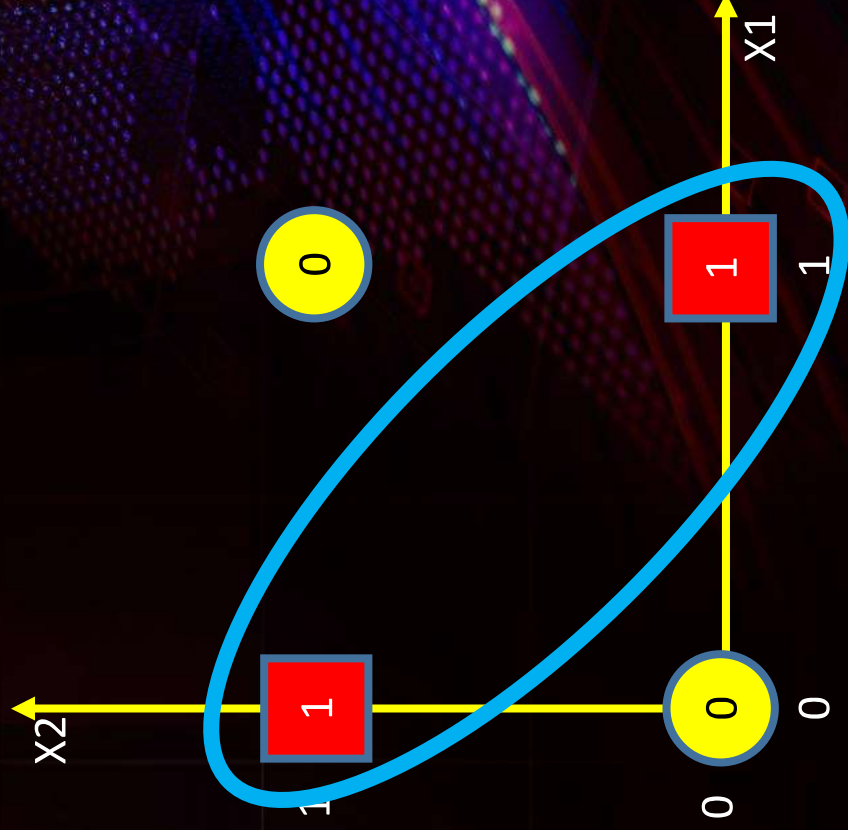


TÓPICOS

1. Limitações dos modelos lineares
2. Redes com múltiplas camadas
3. Interpretação do problema XOR
4. Topologias alternativas

O PROBLEMA XOR

1. Como separar as duas classes?
2. Não é possível com uma fronteira linear
3. Neurônios lineares criam apenas uma fronteira de separação linear

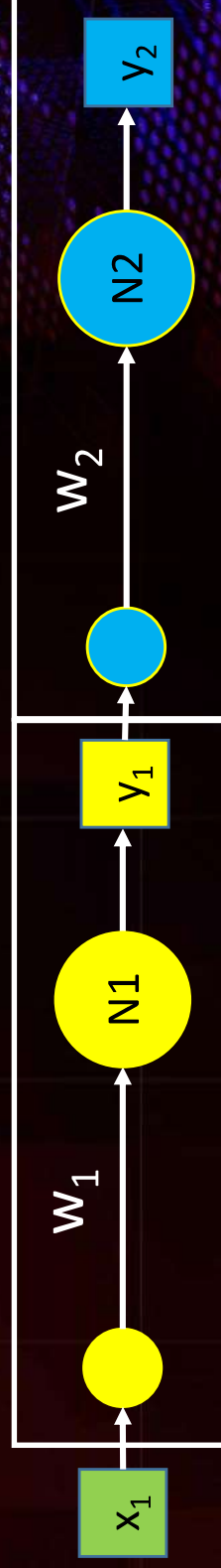


REDES COM MÚLTIPLAS CAMADAS

- Redes com apenas uma camada só representam funções linearmente separáveis
- Redes de múltiplas camadas solucionam essa restrição
- O desenvolvimento do algoritmo de retropropagação foi um dos motivos para o ressurgimento do interesse pela área de redes neurais, na década de 80

REDES COM MÚLTIPLAS CAMADAS

CONTUDO, APENAS SE UTILIZARMOS FUNÇÕES DE ATIVAÇÃO NÃO-LINEARES



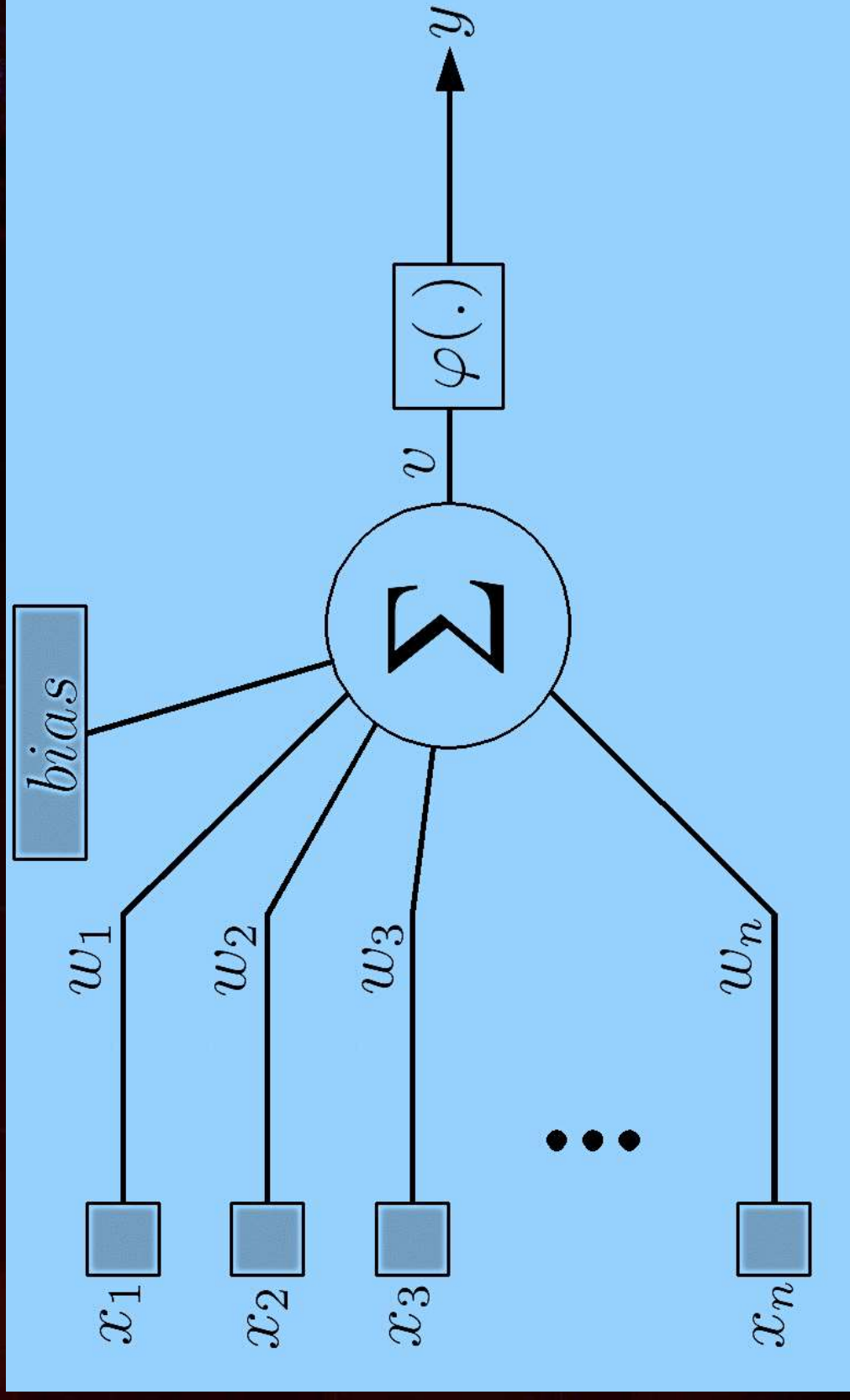
$$y_1 = x_1 w_1$$

$$y_2 = y_1 w_2$$

$$w_c = w_2 w_1$$

$$y_2 = x_1 w_c$$

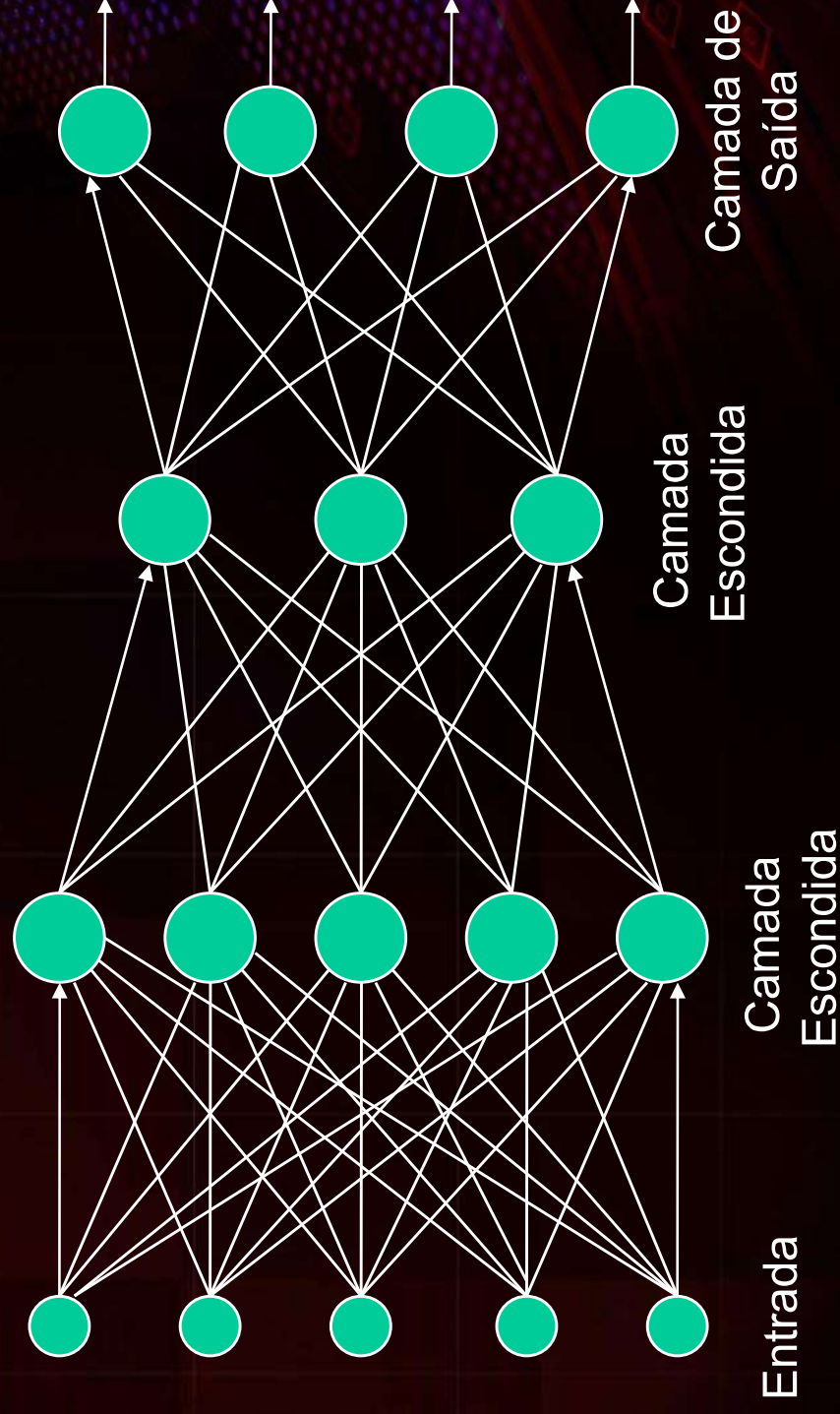
O NEURÔNIO MCP - Revisitando



REDE MLP

REDE COM 3 CAMADAS

- Duas ocultas (escondidas)
- Uma camada de saída



FUNÇÕES DE ATIVAÇÃO

Diversas:

- Linear
- Degrau (Heaviside)
- Sigmoides Logística
- Tangente Hiperbólica
- ReLU (e variações)
- Softmax
- Dentre outras (Ver Cap. 4.1 Livro - Dive Into Deep Learning)

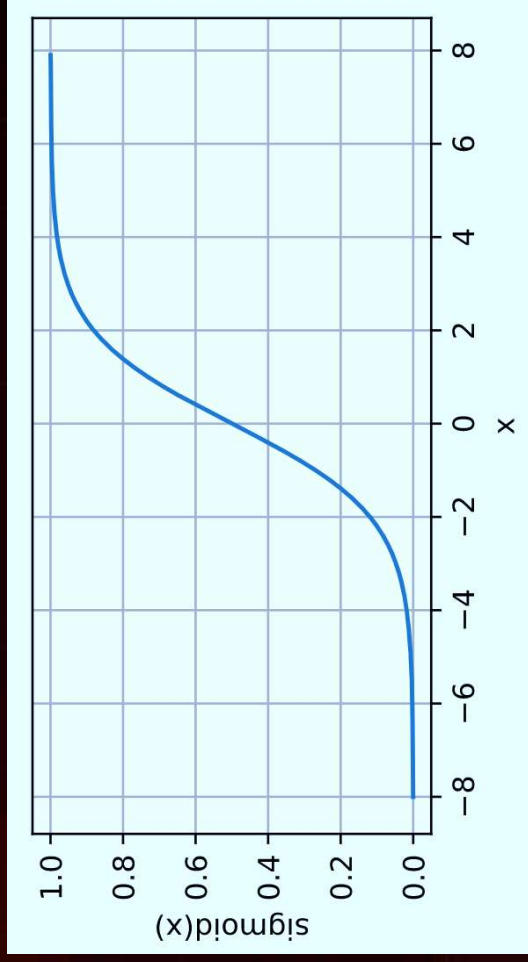
FUNÇÃO LINEAR E DEGRAU

$$f(x) = x$$

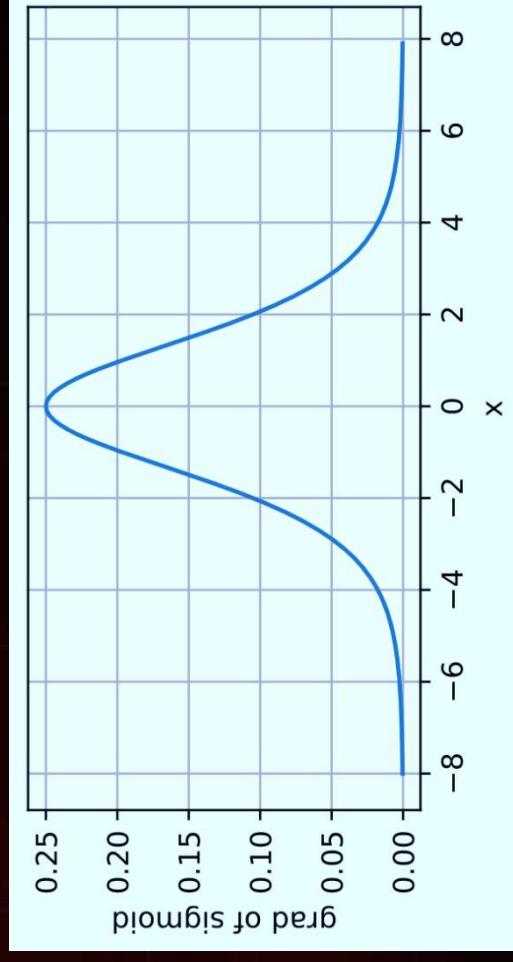
$$f(x) = \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{caso contrario} \end{cases}$$

Fonte: https://pt.d2l.ai/chapter_multilayer-perceptrons/mlp.html

SIGMOIDE LOGÍSTICA

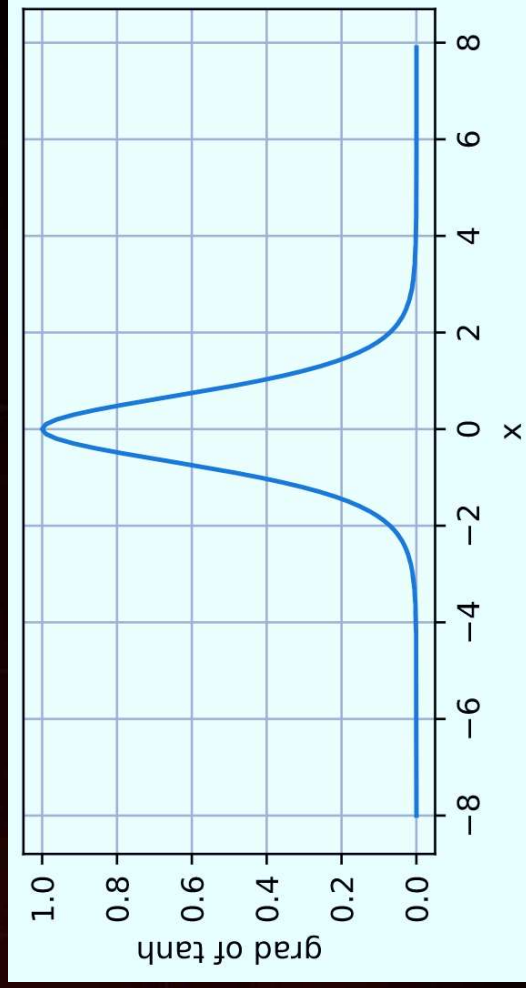
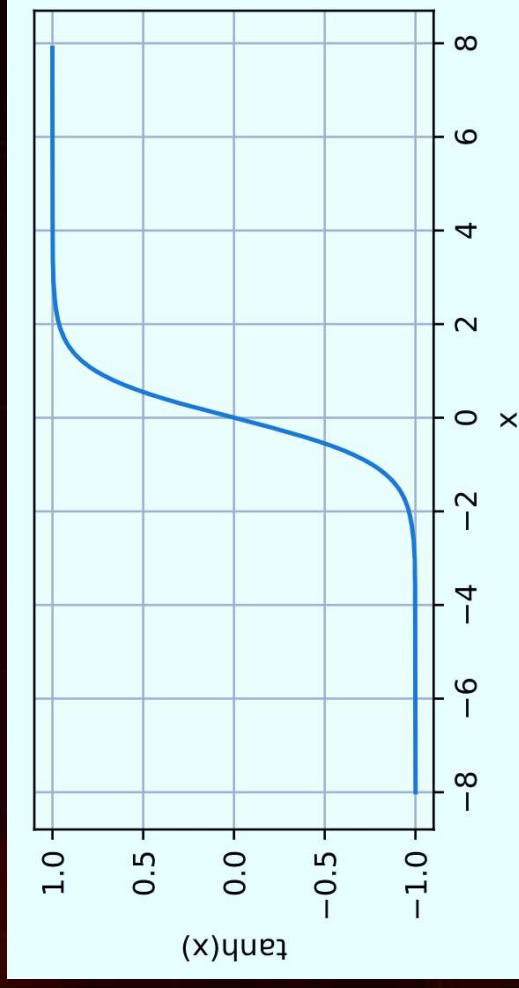


$$f(x) = \frac{1}{1 + \exp(-x)}$$



Fonte:
https://pt.d2l.ai/chapter_multilayer-perceptrons/mlp.html

TANGENTE HIPERBÓLICA

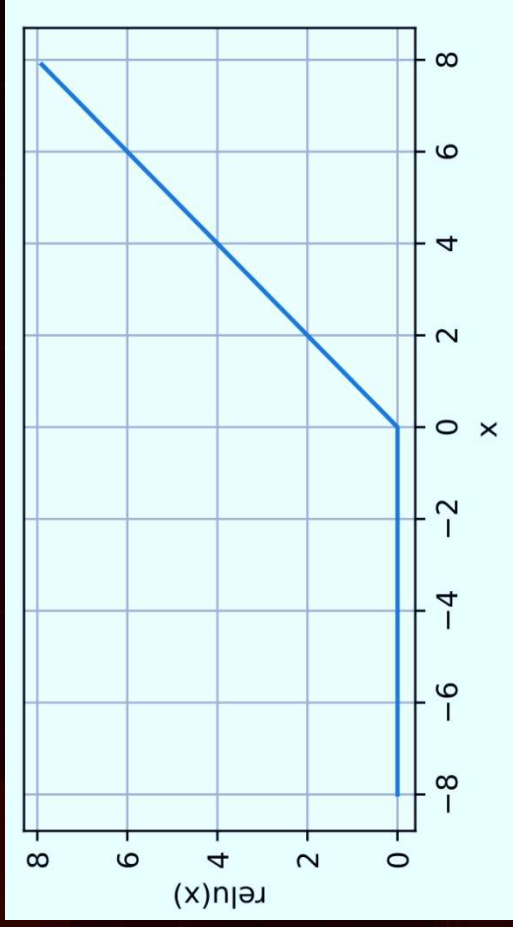


$$f(x) = \tanh(x)$$

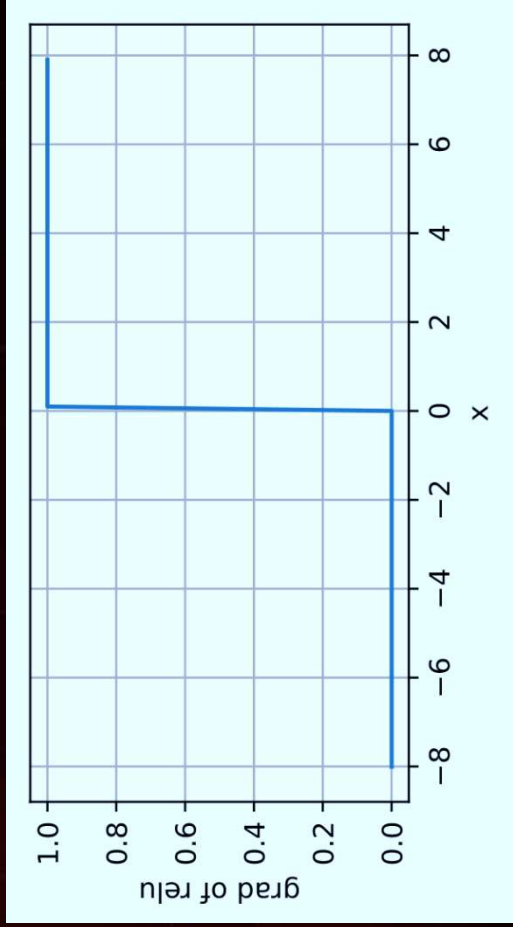
Fonte:

https://pt.d2l.ai/chapter_multilayer-perceptrons/mlp.html

RECTIFIED LINEAR UNIT - ReLU



$$f(x) = \max(0, x)$$



Fonte:

https://pt.d2l.ai/chapter_multilayer-perceptrons/mlp.html

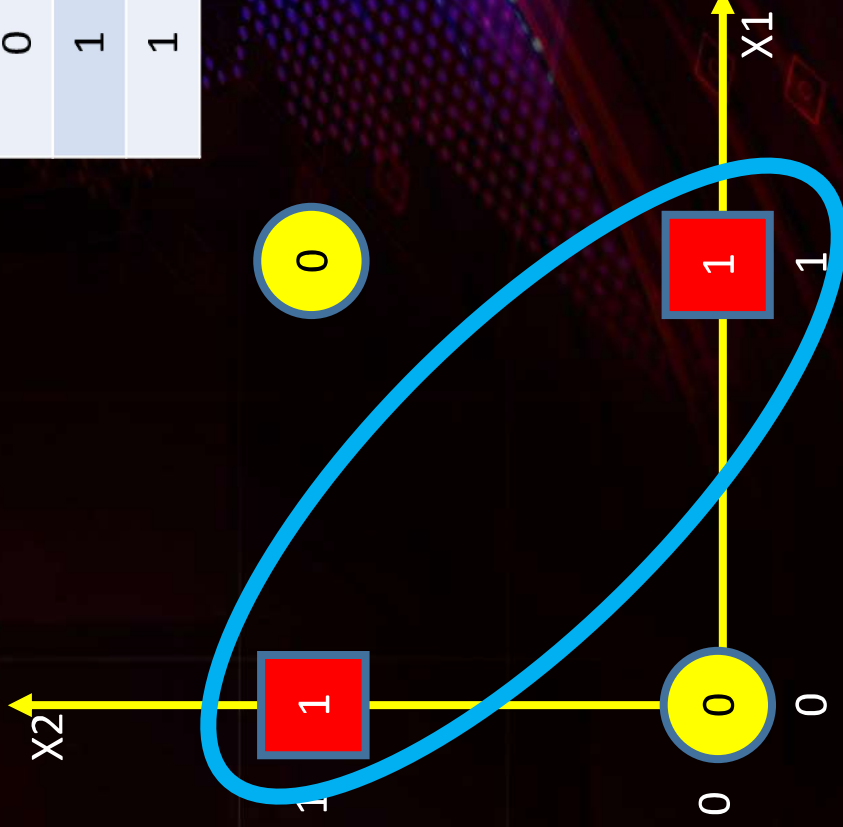
FUNÇÃO SOFTMAX

- Converte a saída padrão em uma distribuição de probabilidade
- Utilizada em problemas de classificação com múltiplas classes

$$f(y) = \frac{\exp(y)}{\sum_k \exp(y_k)}$$

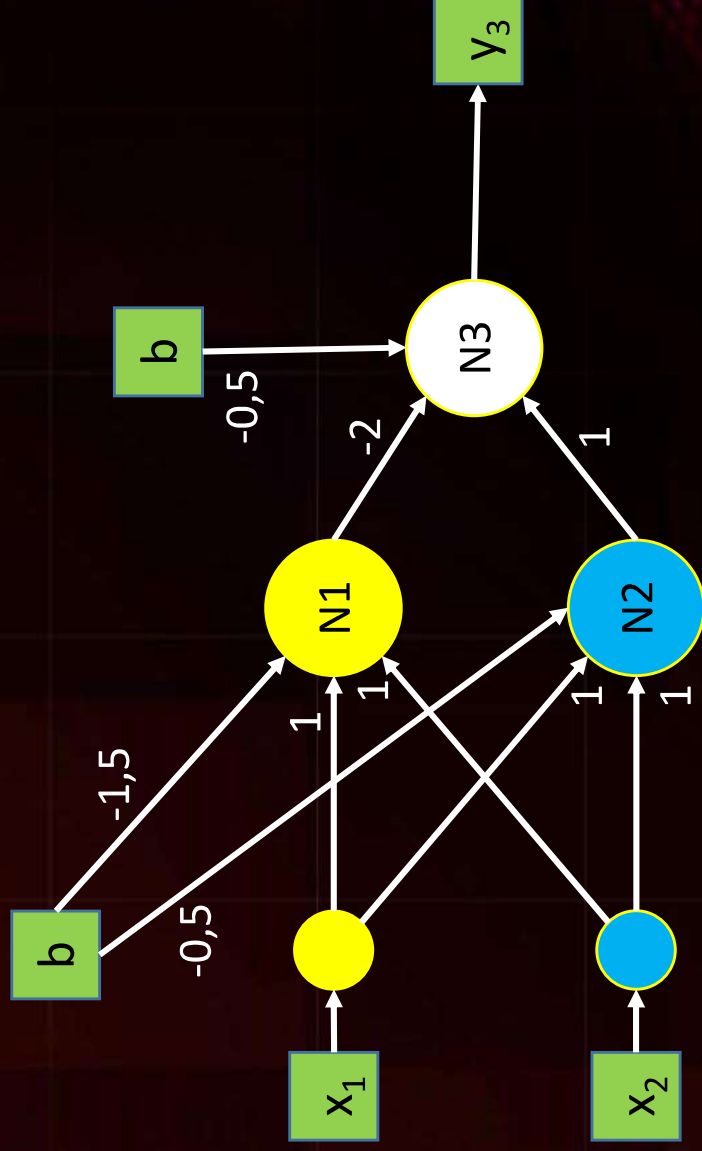
RETOMANDO O PROBLEMA XOR

- Como resolvê-lo com redes de múltiplas camadas?
- Qual é a configuração necessária?



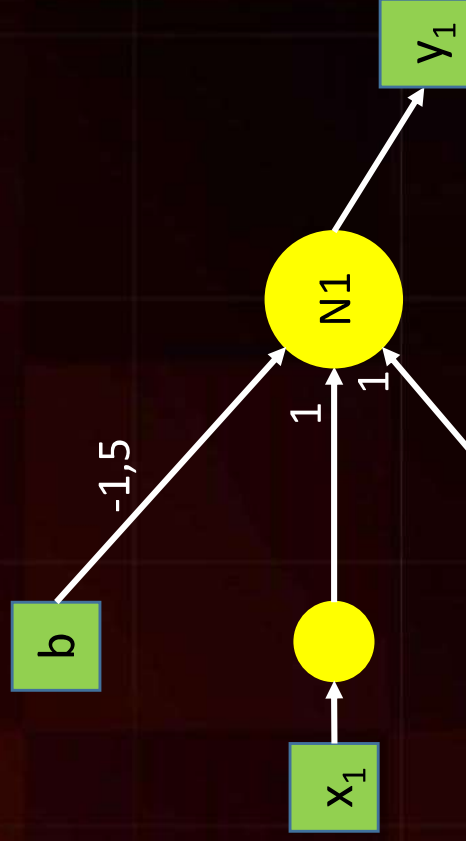
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

PROBLEMA XOR



x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

PROBLEMA XOR



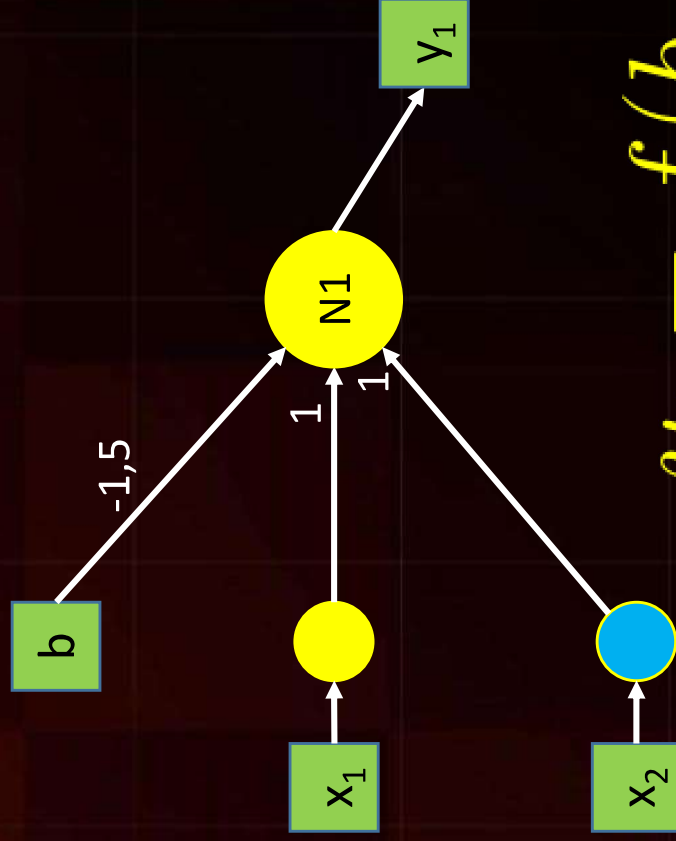
x_1	x_2	y_1	y_3
0	0	?	0
0	1	?	1
1	0	?	1
1	1	?	0

$$y_1 = f(b + x_1w_1 + x_2w_2)$$

$$y_1 = f(-1,5 + 0 + 0)$$

$$y_1 = 0$$

PROBLEMA XOR



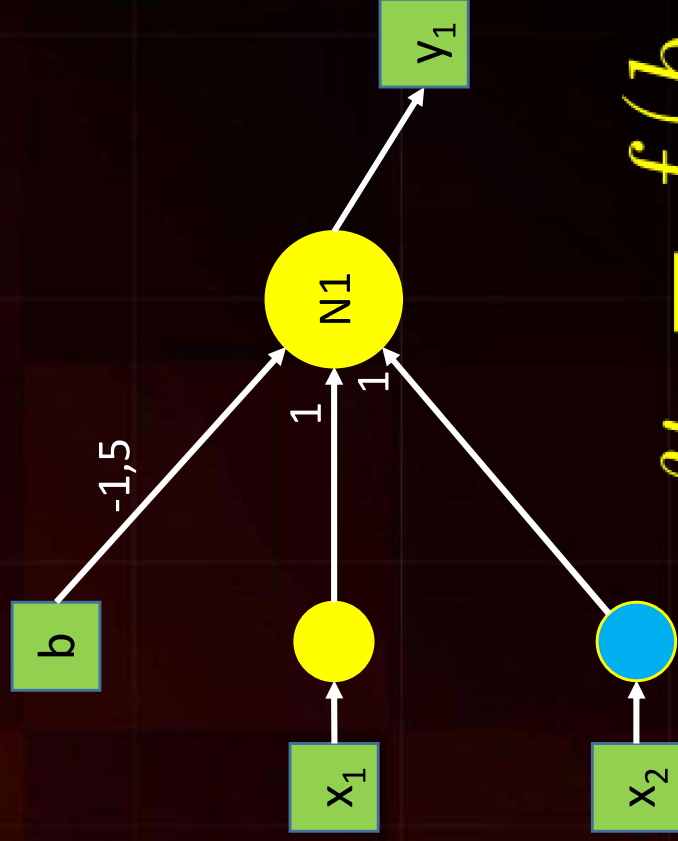
x_1	x_2	y_1	y_3
0	0	0	0
0	1	?	1
1	0	?	1
1	1	?	0

$$y_1 = f(b + x_1w_1 + x_2w_2)$$

$$y_1 = f(-1.5 + 0 + 1 * 1)$$

$$y_1 = 0$$

PROBLEMA XOR



x_1	x_2	y_1	y_3
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

$$y_1 = f(b + x_1w_1 + x_2w_2)$$

$$y_1 = f(-1,5 + 1 + 1)$$

$$y_1 = 1$$

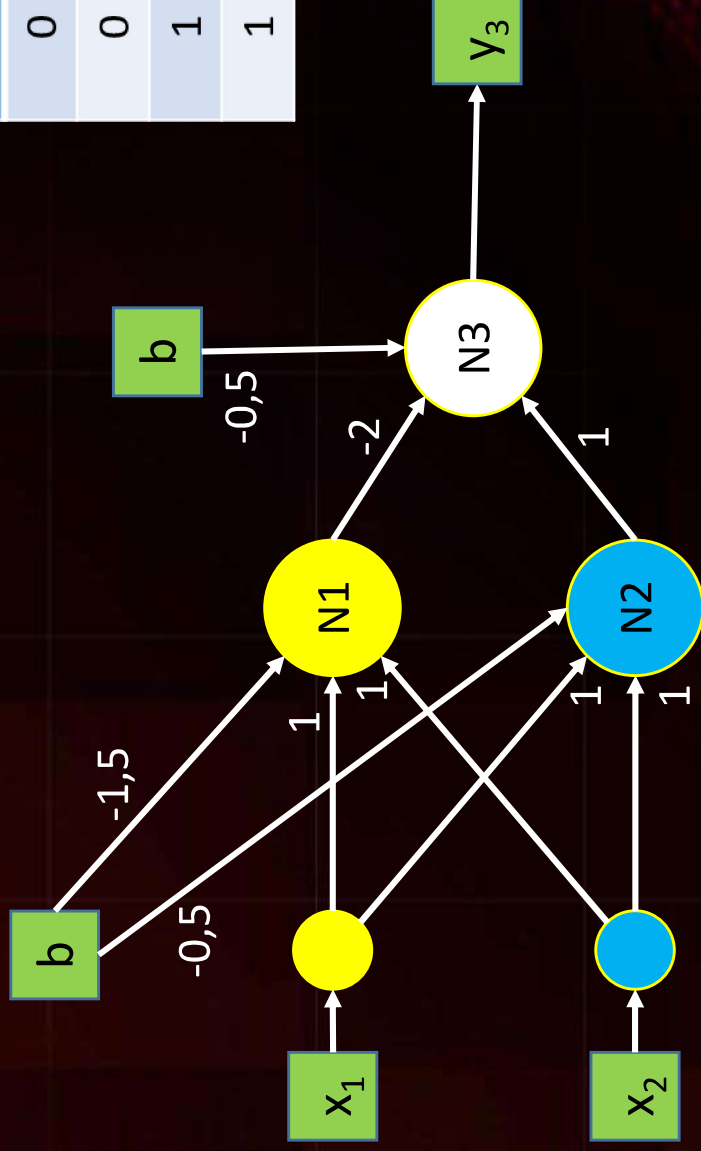
PROBLEMA XOR

- O que o neurônio N1 representa?

x_1	x_2	y_1	y_3
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

- A saída y_1 é 1 apenas quando as duas entradas estiverem ativas ($=1$)
- **Porta Lógica AND (E-Lógico)**

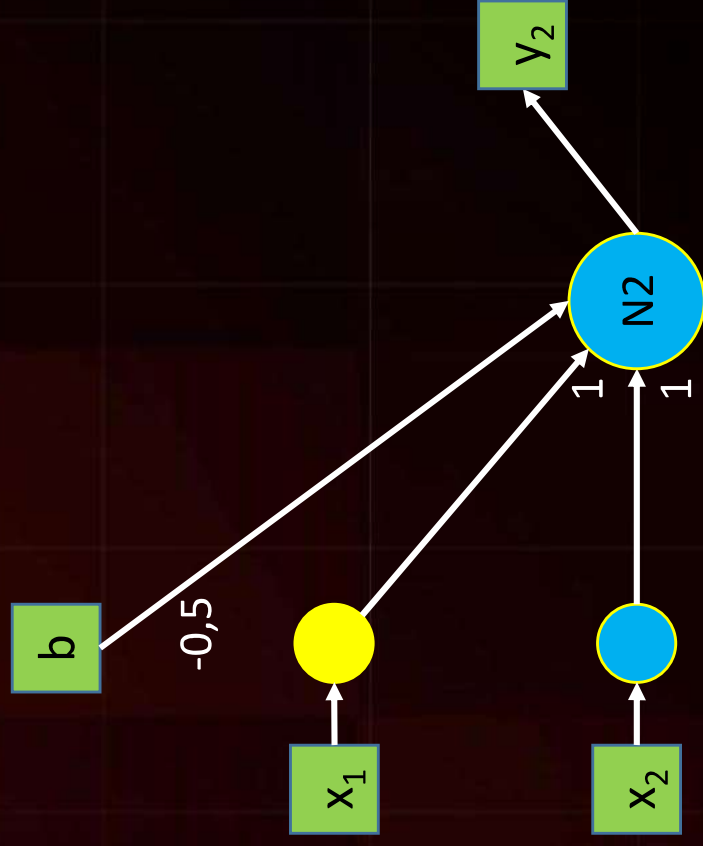
PROBLEMA XOR



x_1	x_2	y_1	y_2	y_3
0	0	0	?	0
0	1	0	?	1
1	0	0	?	1
1	1	1	?	0

O que o Neurônio N2 representa?

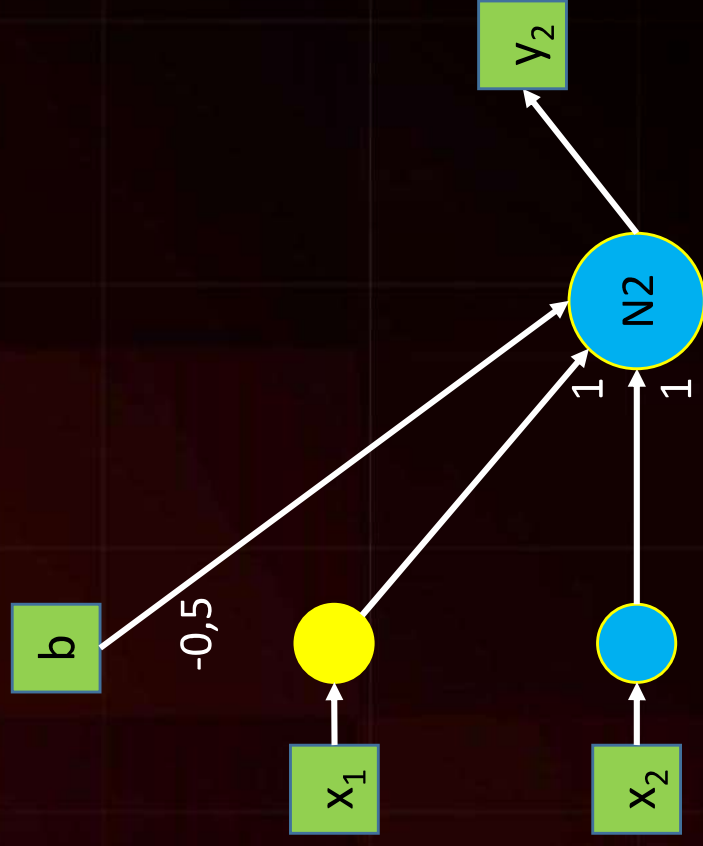
PROBLEMA XOR



x_1	x_2	y_1	y_2	y_3
0	0	0	?	0
0	1	0	?	1
1	0	0	?	1
1	1	1	?	0

$$y_2 = f(b + x_1w_1 + x_2w_2)$$

PROBLEMA XOR

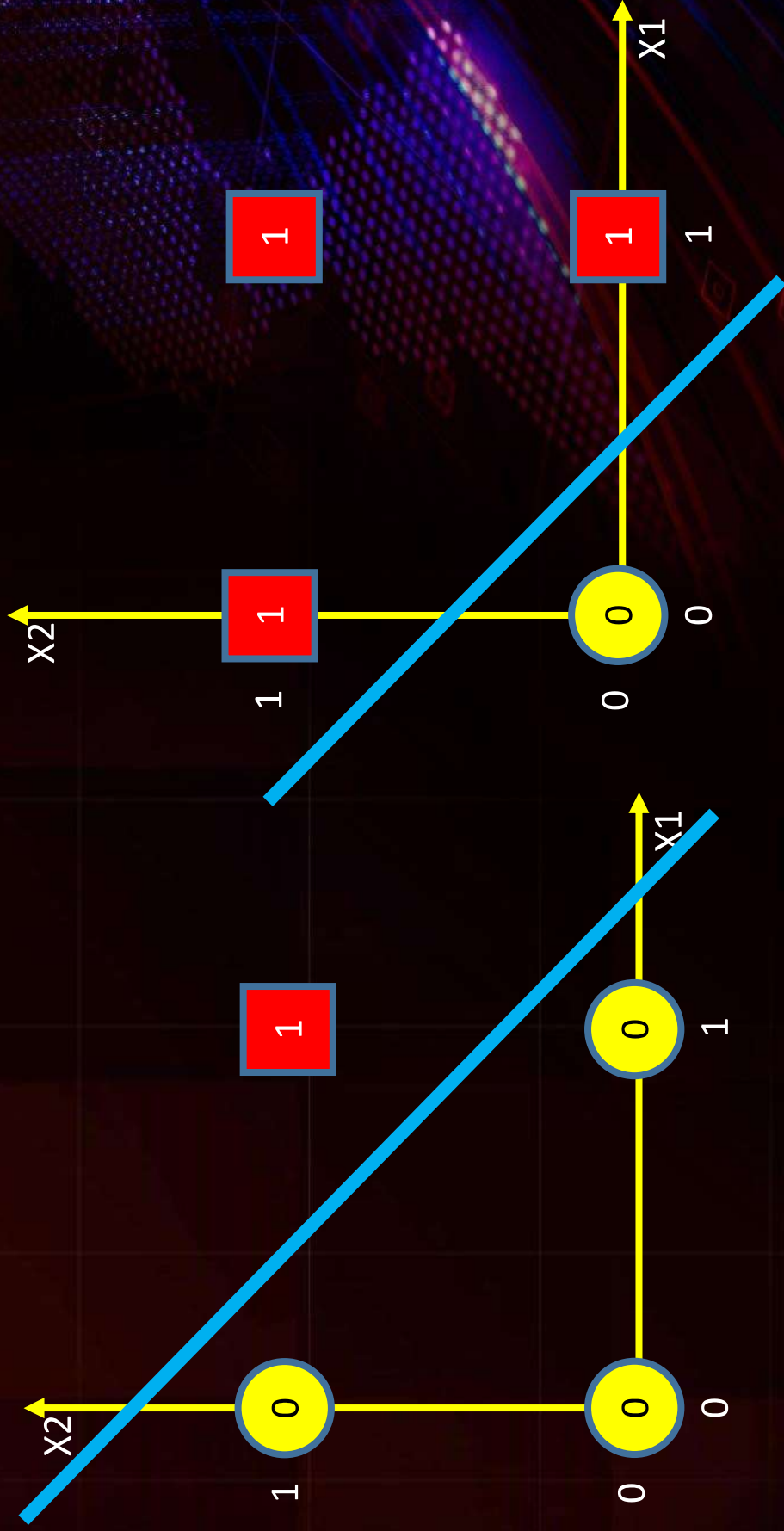


Porta Lógica OR

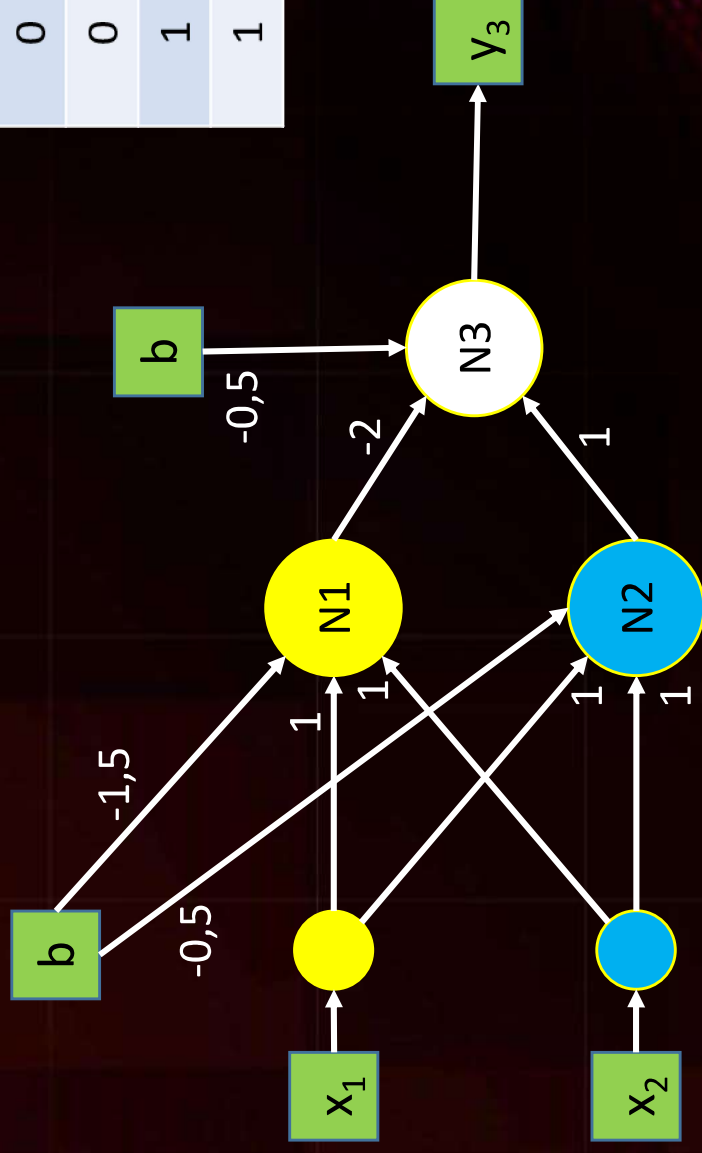
x_1	x_2	y_1	y_2	y_3
0	0	0	0	0
0	1	0	1	1
1	0	0	1	1
1	1	1	1	0

$$y_2 = f(b + x_1w_1 + x_2w_2)$$

PROBLEMA XOR



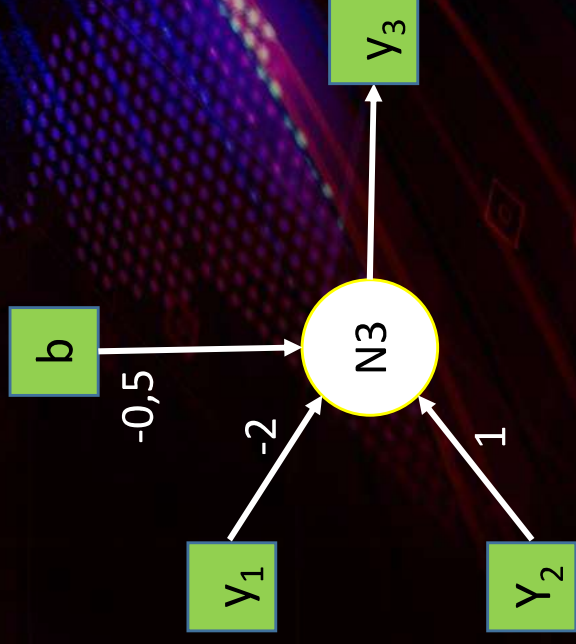
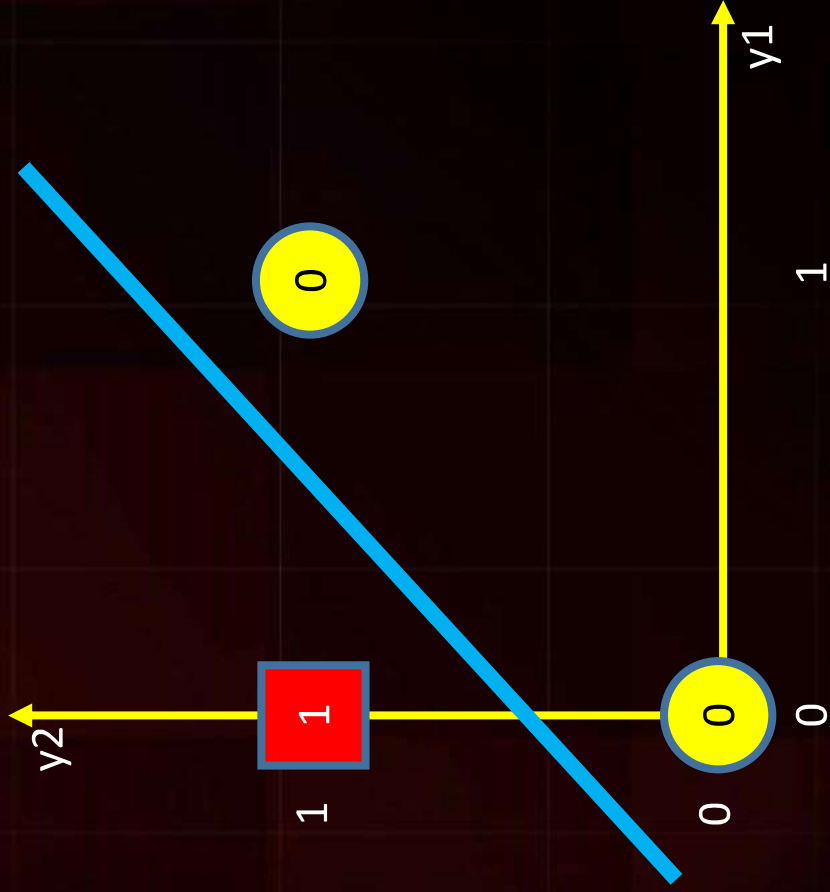
PROBLEMA XOR



x_1	x_2	y_1	y_2	y_3
0	0	0	0	0
0	1	0	1	1
1	0	0	1	1
1	1	1	1	0

O que o Neurônio $N3$ representa?

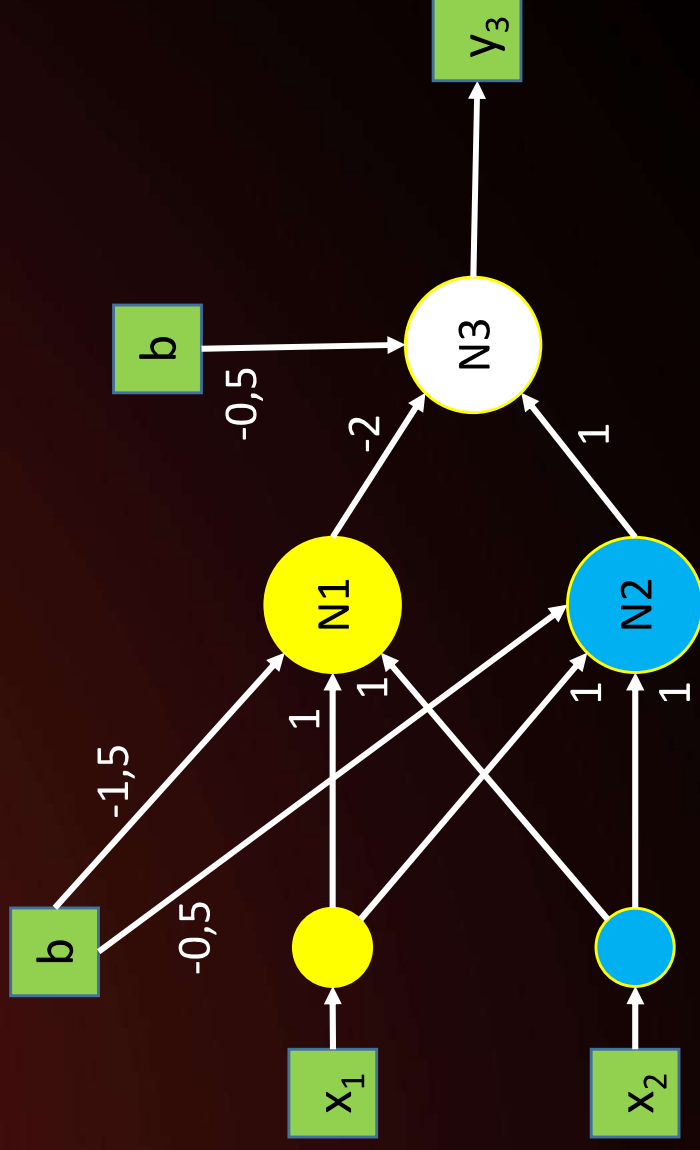
PROBLEMA XOR



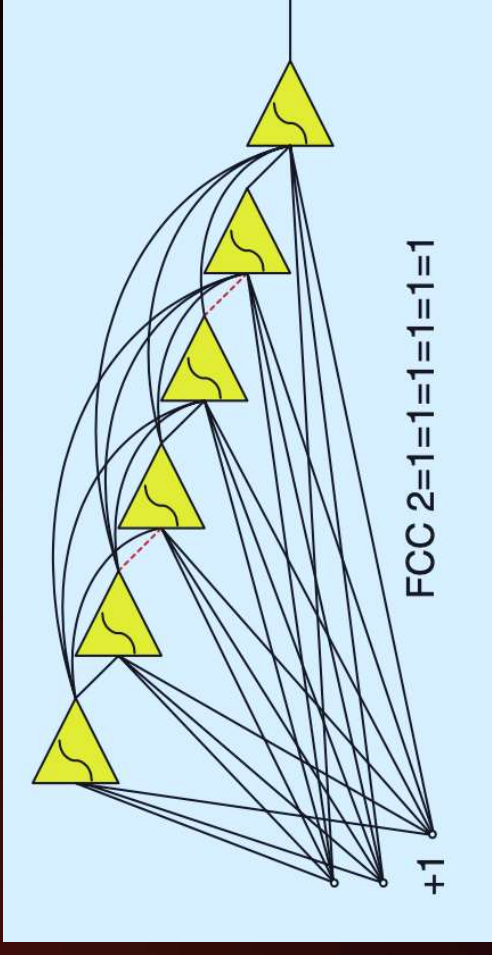
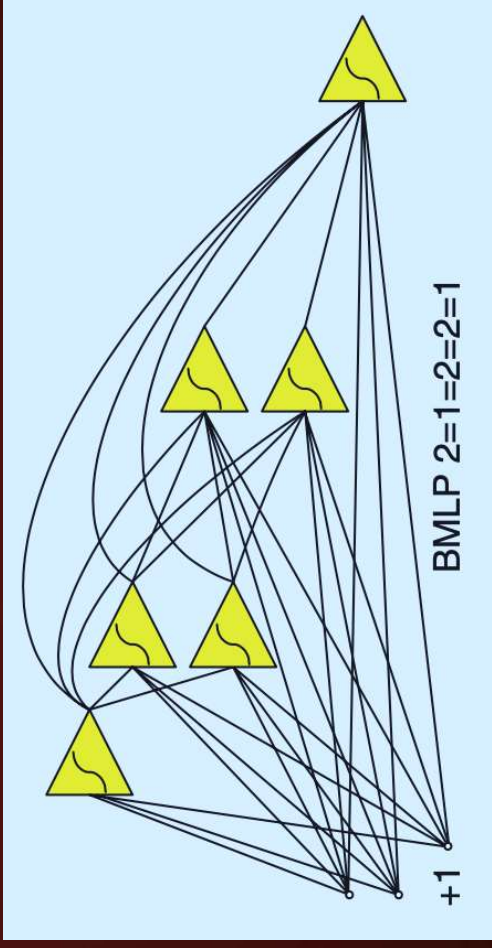
y_1	y_2	y_3
0	0	0
0	1	1
0	1	1
1	1	0

O QUE A REDE MLP FAZ?

Resolve problemas não-linearmente separáveis a partir da transformação do problema original em um problema linearmente separável (camada a camada)



TOPOLOGIAS ALTERNATIVAS



Fonte: Wilamowski (2009), IEEE Industrial Electronics Magazine
Disponível em: <https://www.eng.auburn.edu/~wilambm/pap/>

O QUE VIMOS?

- Entendemos a limitação das redes com única camada e com múltiplas camadas lineares
- Conhecemos a rede MLP e a função dos neurônios das camadas ocultas
- Vimos como a rede MLP resolve o problema XOR

PRÓXIMA VIDEOAULA

- Aprenderemos a treinar uma rede MLP com o algoritmo de retropropagação (backpropagation)
- Entenderemos a derivação do algoritmo de retropropagação

ATÉ A PRÓXIMA!