

REDES NEURAIS

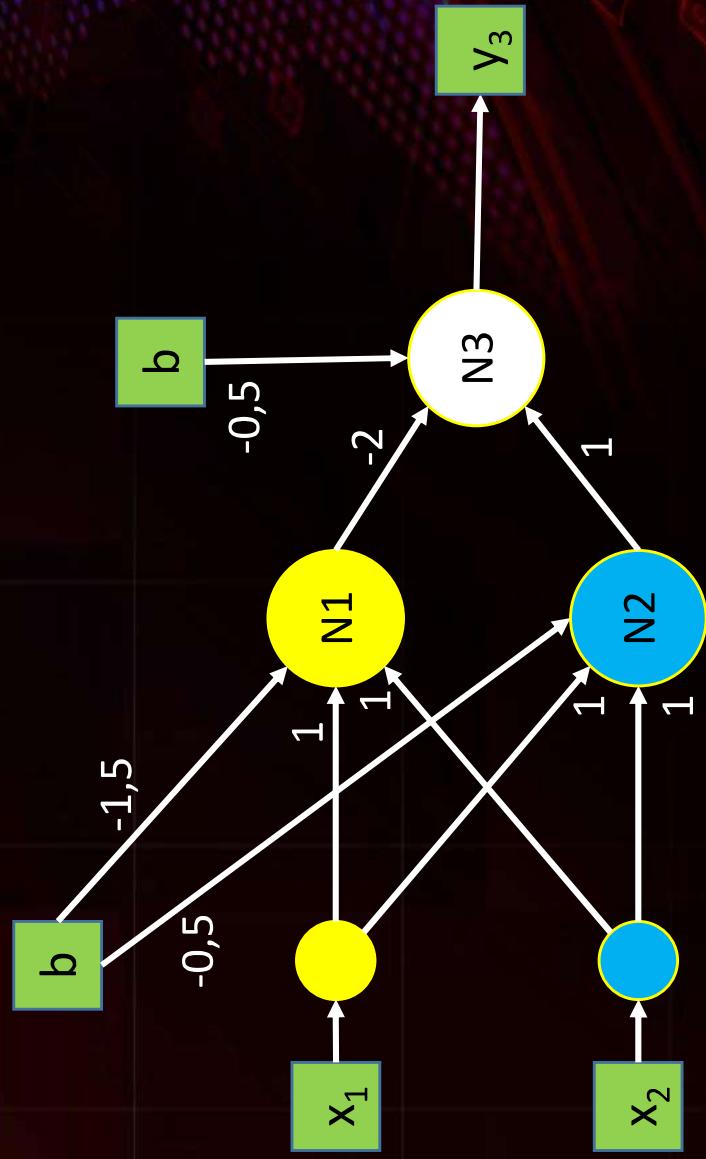
**Fundamentos e derivação do
algoritmo de retropropagação**

TÓPICOS

1. Introdução ao algoritmo de Retropropagação
2. Derivação do Algoritmo de Retropropagação

RETRÔPROPAGAÇÃO

PROBLEMA CENTRAL: como atualizar os neurônios das camadas ocultas (escondida) da rede?



RETRORROPOAGAÇÃO

- O algoritmo de retropropagação surgiu para resolver esse problema
- O seu desenvolvimento mostrou que redes com múltiplas camadas poderiam ser treinadas e resolver problemas não-lineares
- Motivou o interesse pela área na década de 80

RETRÔPROPAGAÇÃO

Processo de minimização do erro pelo método do gradiente descendente:

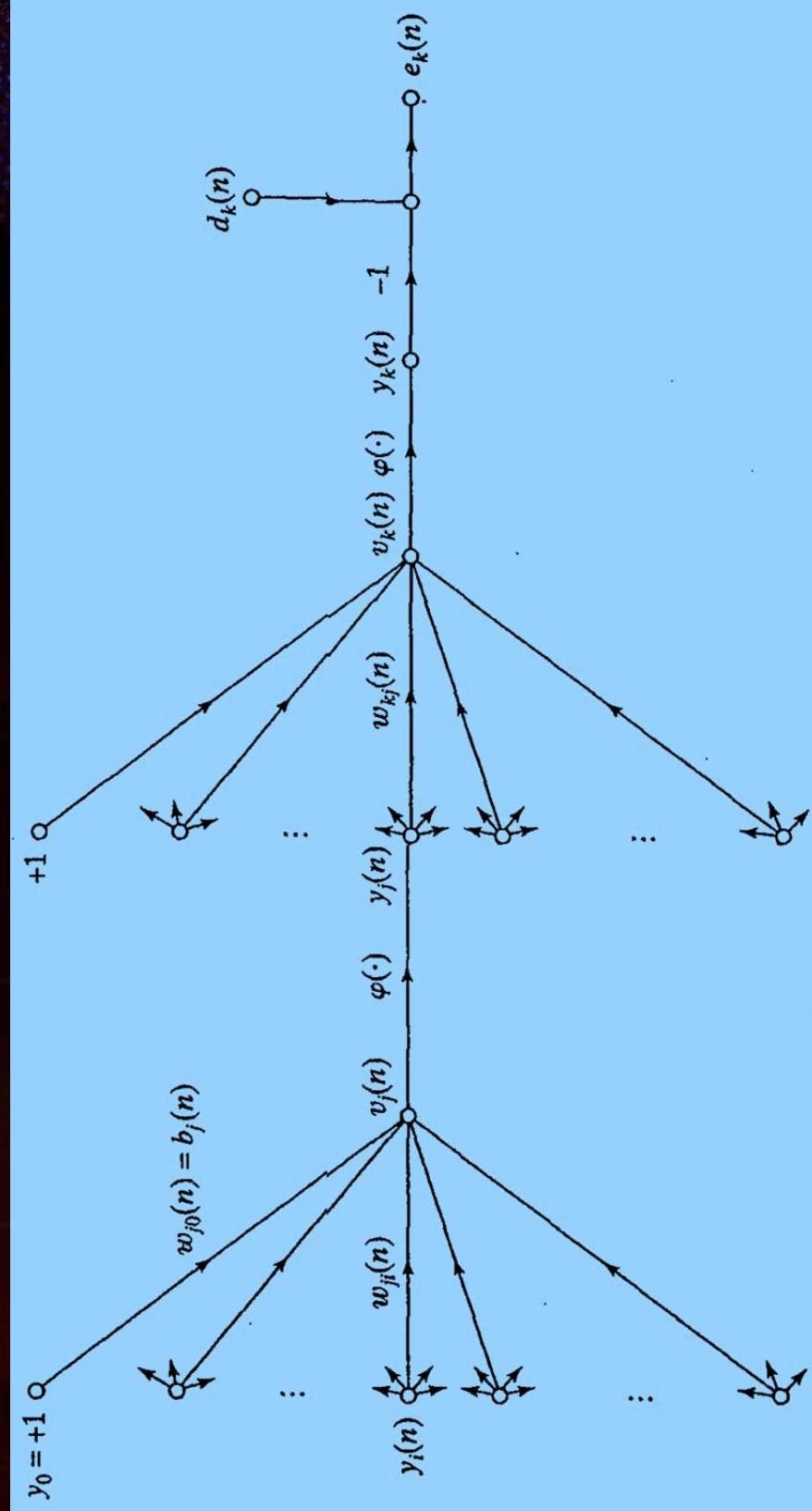
$$\Delta w_{ki} = -\eta \frac{\partial E}{\partial w_{ki}}$$

$$E(n) = \frac{1}{2} \sum (d_k(n) - y_k(n))^2$$

RETRÔPROPAGAÇÃO

GRAFO DE FLUXO:

$$\Delta w_{ki} = -\eta$$



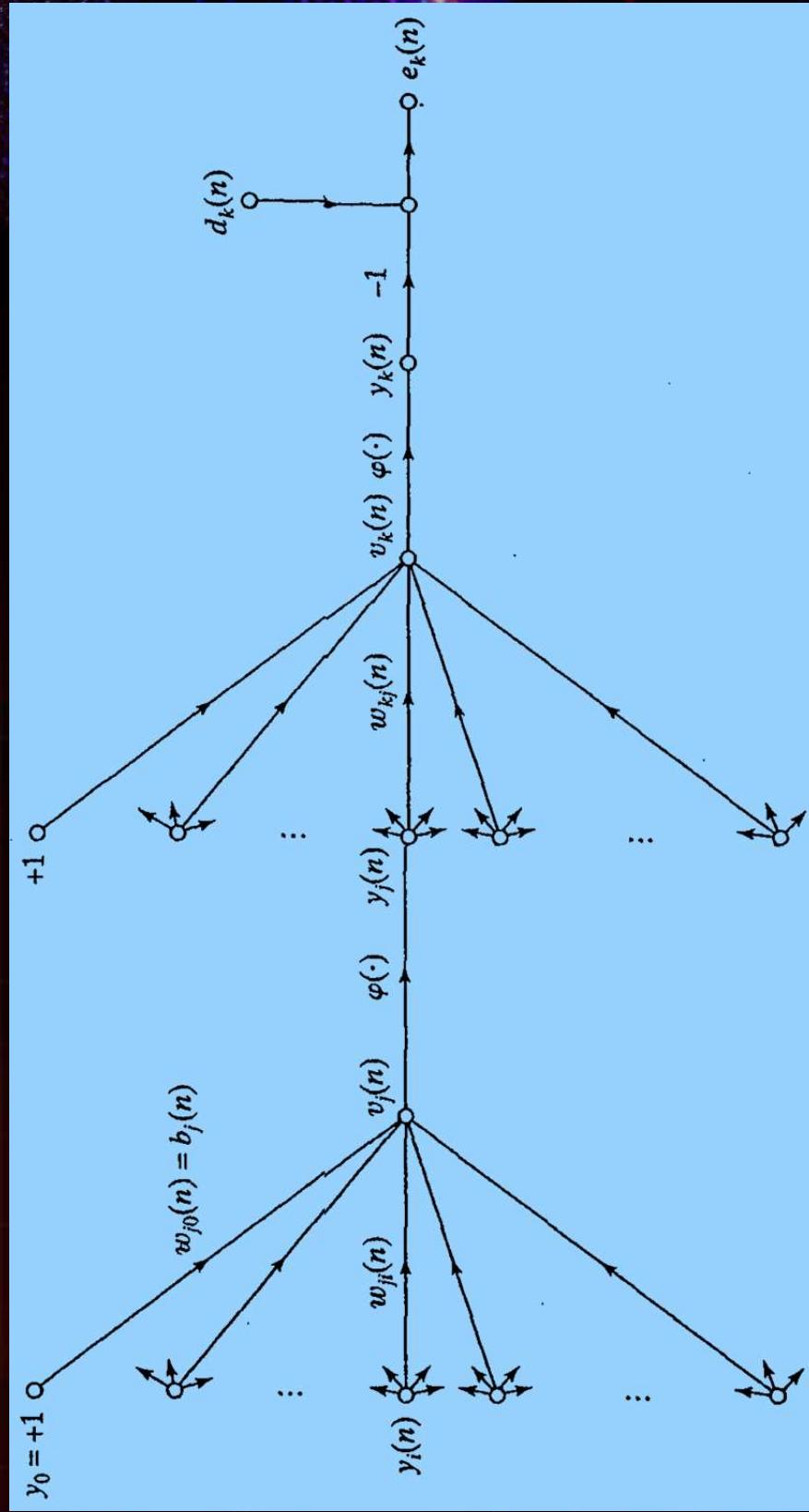
RETRÔPROPAGAÇÃO

$$\Delta w_{kj} = -\eta \frac{\partial E}{\partial w_{kj}} = -\eta \frac{\partial E}{\partial v_k} \frac{\partial v_k}{\partial w_{kj}}$$
$$v_k = \sum w_{kj} x_j$$
$$\delta_k = -\frac{\partial E}{\partial v_k}$$
$$\Delta w_{kj} = \eta \delta_k x_j$$

GRADIENTE LOCAL

δ_k

COMO CALCULAR O GRADIENTE LOCAL?



GRADIENTE LOCAL

$$\delta_k = -\frac{\partial E}{\partial v_k} = -\frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial v_k}$$

$$\boxed{\frac{\partial E}{\partial y_k}}$$

$$y_k = f(v_k)$$

Depende da posição
do neurônio na rede

CASO 1: NEURÔNIO DE SÁIDA

$$\delta_k = -\frac{\partial E}{\partial v_k} = -\boxed{\frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial v_k}}$$

$$E(n) = \frac{1}{2} \sum (d_k(n) - y_k(n))^2$$
$$2 \frac{1}{2} (d_k - y_k)(-1) = -e_k$$

$$y_k = f(v_k)$$
$$f'(v_k)$$

$$\delta_k = e_k f'(v_k)$$

CASO 2: NEURÔNIO OCULTO

$$\delta_j = -\frac{\partial E}{\partial v_j} = -\frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial v_j}$$

$$y_j = f(v_j)$$

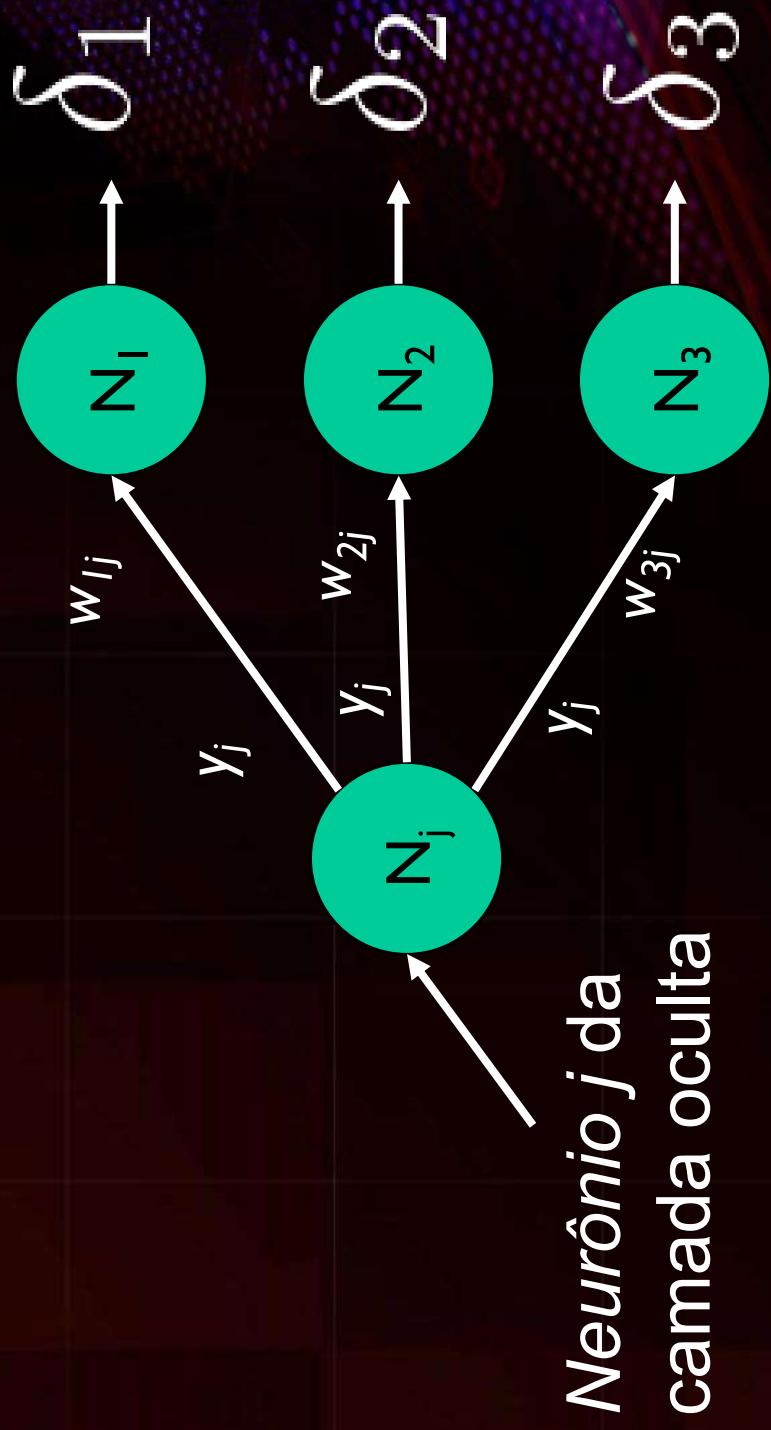
Qual é o erro do
neurônio oculto?

FUNÇÕES DE ATIVAÇÃO

PELO APRENDIZADO SUPERVISIONADO, SÓ CONHECEMOS O ERRO NA SAÍDA DO MODELO

- O erro no neurônio k é uma função do estado interno do neurônio k
- O estado interno do neurônio k depende das suas entradas
- Portanto, um neurônio j da camada anterior pode afetar o estado do neurônio k em maior ou menor grau

CASO 2: NEURÔNIO OCULTO



CASO 2: NEURÔNIO OCULTO

$$\delta_j = -\frac{\partial E}{\partial v_j} = -\frac{\partial E}{\partial v_j} \frac{\partial y_j}{\partial v_j}$$

$y_j = f(v_j)$

$f'(v_j)$

$$E = \frac{1}{2} \sum_k e_k^2$$

$\sum_k e_k \frac{\partial e_k}{\partial y_j}$

$$\delta_j = -f'(v_j) \sum_k e_k \frac{\partial e_k}{\partial y_j}$$

CASO 2: NEURÔNIO OCULTO

$$\delta_j = -f'(v_j) \sum_k e_k \frac{\partial e_k}{\partial y_j}$$

$$\delta_j = -f'(v_j) \sum_k e_k \frac{\partial e_k}{\partial v_k} \frac{\partial v_k}{\partial y_j}$$

$$e_k = d_k - f(v_k)$$

$$\frac{\partial e_k}{\partial v_k}$$

$$\frac{\partial v_k}{\partial y_j} = w_{kj}$$

$$v_k = \sum_j w_{kj}$$

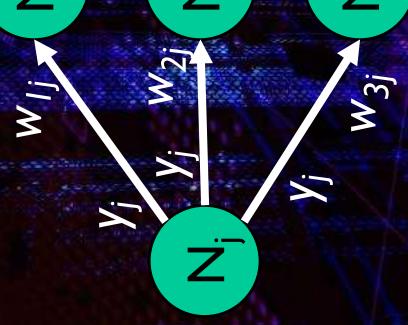
CASO 2: NEURÔNIO OCULTO

Logo, juntando tudo temos:

$$\delta_j = f'(v_j) \sum_k e_k f'(v_k) w_{kj}$$

Ou ainda:

$$\delta_j = f'(v_j) \sum_k \delta_k w_{kj}$$



ATUALIZAÇÃO

$$\text{Temos: } \Delta w_{kj} = \eta \delta_k x_j$$

Para k sendo um neurônio de saída:

$$\Delta w_{kj} = \eta f'(v_k) e_k x_j$$

Para os neurônios ocultos:

$$\Delta w_{ji} = \eta f'(v_j) \left(\sum_k w_{kj} \delta_k x_i \right)$$

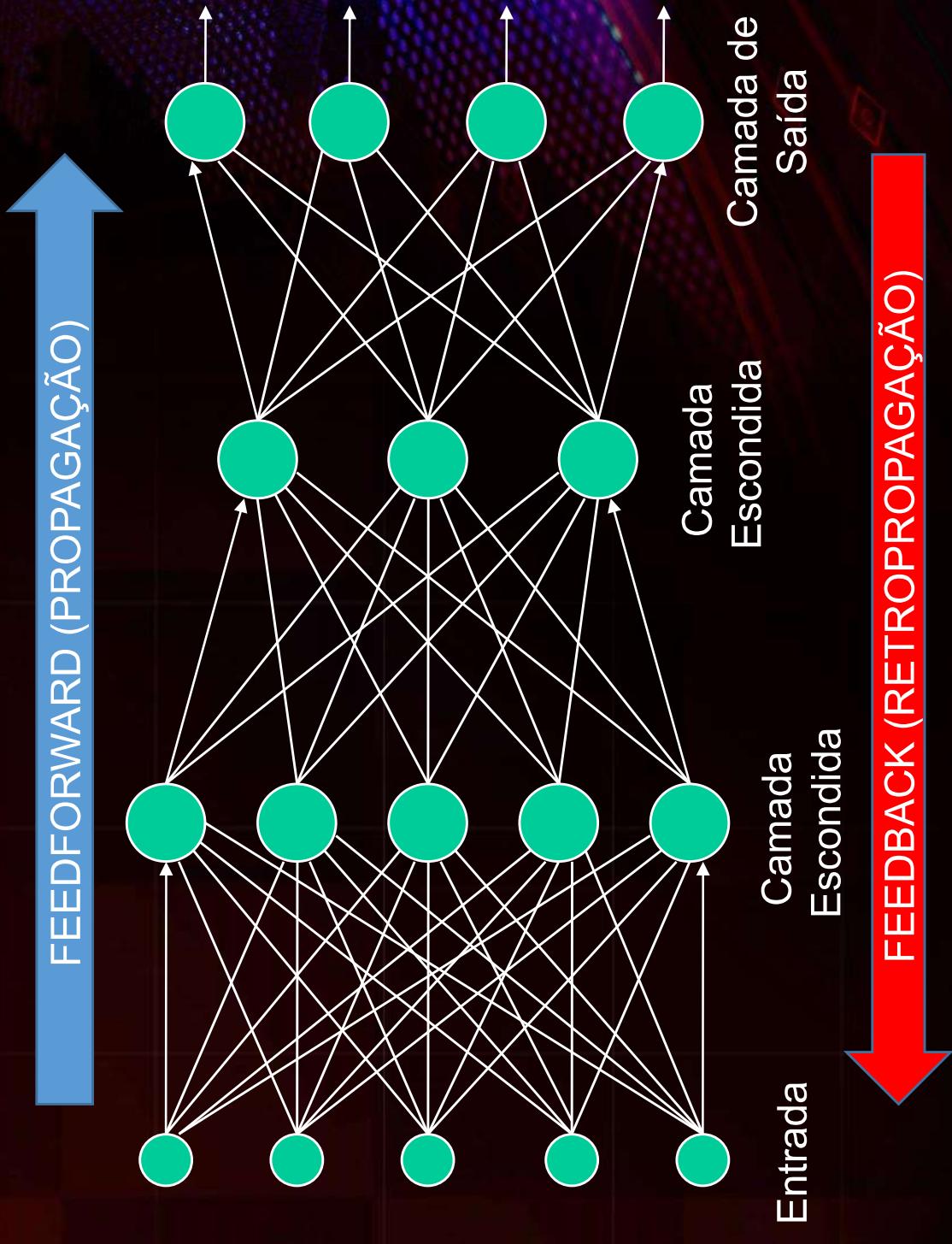
A RETROPROPAGAÇÃO

O ALGORITMO DE RETROPROPAGAÇÃO POSSUI DUAS FASES:

- Feedforward (adiante): as entradas se propagam pela rede, camada a camada, da entrada até a saída
- Feedback (retorno): os erros se propagam, camada a camada, da saída até a entrada

O algoritmo de retropropagação permite o cálculo do gradiente dos neurônios ocultos

A RETROPROPAGAÇÃO



O ALGORITMO

- Inicialização: hiperparâmetros e pesos
- Treinamento: repita até que o erro seja \leq tolerância para todos os padrões de conjunto de treinamento (ou número máximo de épocas)
 1. Aplica-se um padrão de entrada x_i e seu respectivo vetor de saída d_i desejado;
 2. Propagação do sinal entrada → saída
 3. Retropropagação dos erros saída → entrada
 4. Atualização dos pesos de cada processador
 5. Volta ao passo 1

O QUE VIMOS?

- Compreendemos os fundamentos do algoritmo de retropropagação
- Conhecemos a sua derivação
-

PRÓXIMA VIDEOAULA

- Ilustração prática da rede MLP

ATÉ A PRÓXIMA !!