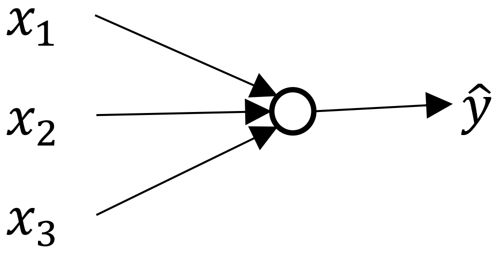
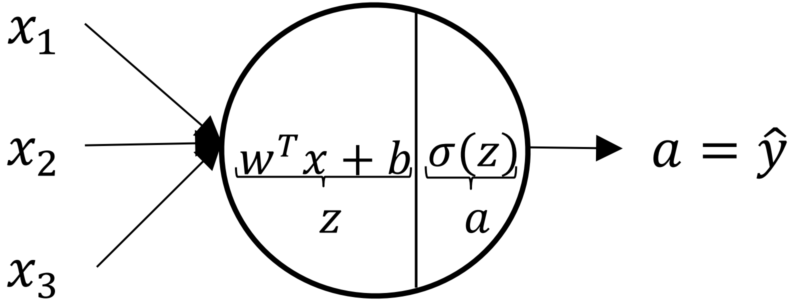
Processamento de Linguagem Natural

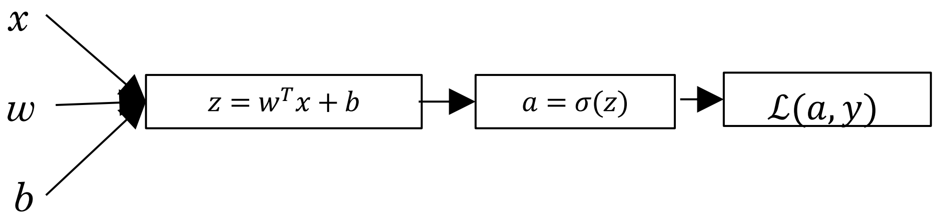
Redes Neurais Multicamadas

Prof. Leandro B. Marinho
lbmarinho@dsc.ufcg.edu.br

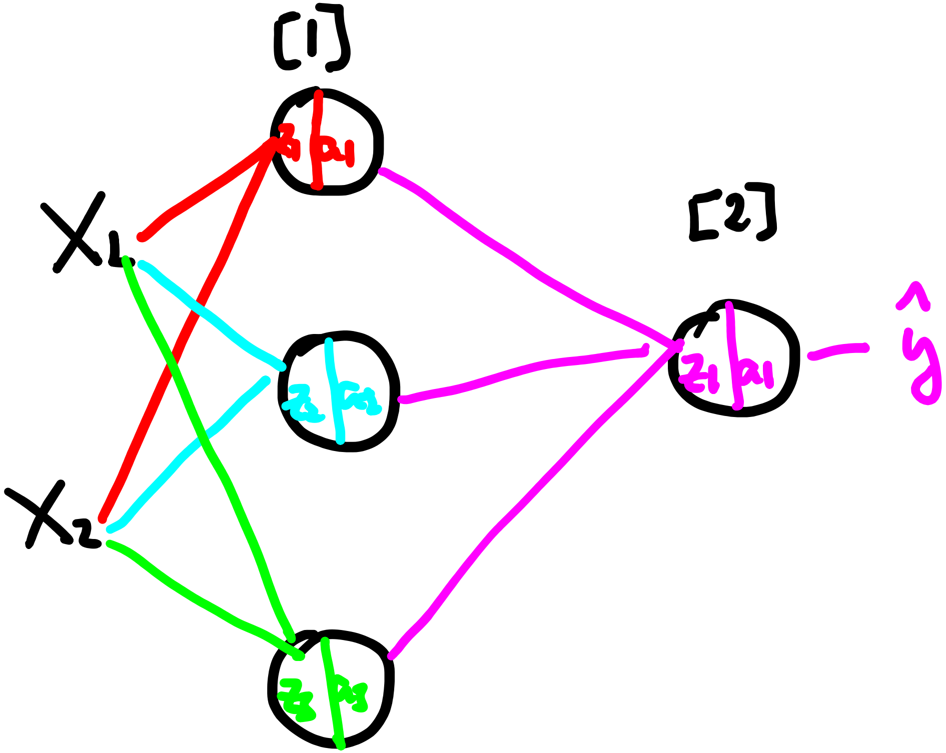
O que é uma Rede Neural?



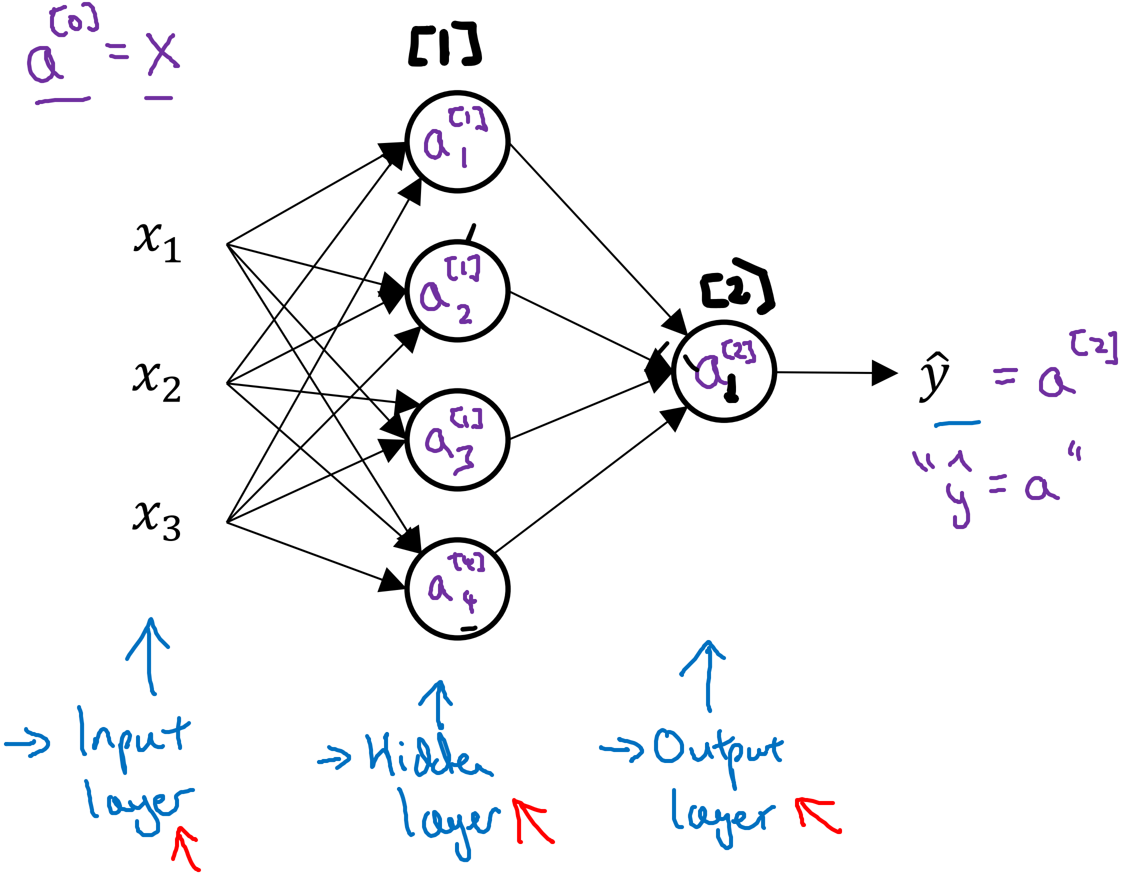




Representação de Redes Neurais

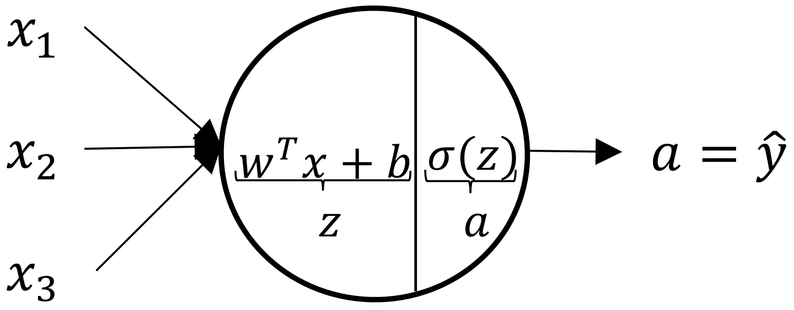


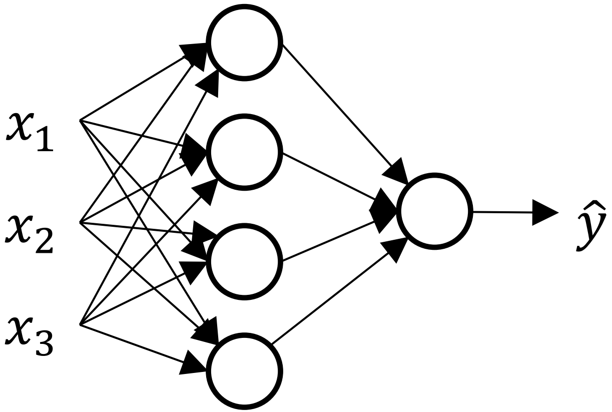
Representação de Redes Neurais



$$a^{[1]} = \begin{bmatrix} a^{[1]}_1 \\ a^{[1]}_2 \\ \vdots \\ a^{[1]}_4 \end{bmatrix}$$

Representação de Redes Neurais

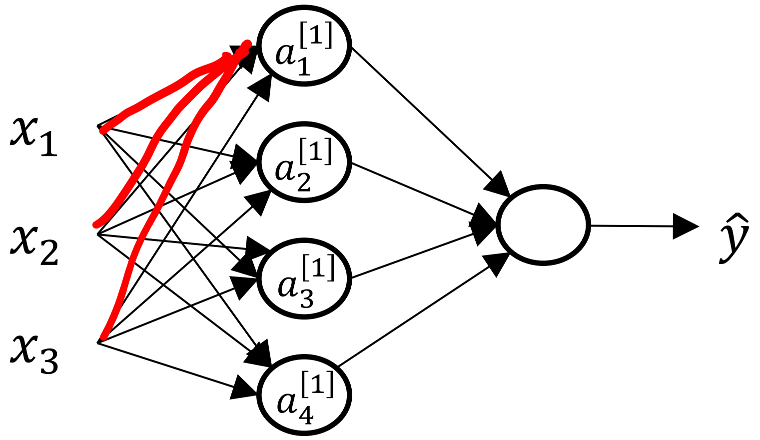




$$z = w^T x + b$$

$$a = \sigma(z)$$

Representação de Redes Neurais



$$z_1^{[1]} = w_1^{[1]T} x + b_1^{[1]},\ a_1^{[1]} = \sigma(z_1^{[1]})$$

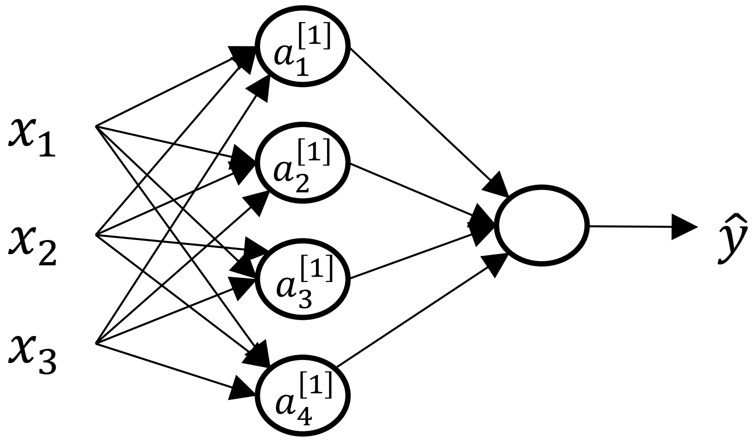
$$z_2^{[1]} = w_2^{[1]T} x + b_2^{[1]},\ a_2^{[1]} = \sigma(z_2^{[1]})$$

$$z_3^{[1]} = w_3^{[1]T} x + b_3^{[1]},\ a_3^{[1]} = \sigma(z_3^{[1]})$$

$$z_4^{[1]} = w_4^{[1]T} x + b_4^{[1]},\ a_4^{[1]} = \sigma(z_4^{[1]})$$

$$z_1^{[2]} = w_1^{[2]T} \cdot a^{[1]} + b_1^{[2]},\ a_1^{[2]} = \sigma(z_1^{[2]})$$

Representação de Redes Neurais



Dada entrada x:

$$z^{[1]} = W^{[1]}x + b^{[1]}$$

$$a^{[1]} = \sigma(z^{[1]})$$

$$z^{[2]} = W^{[2]}a^{[1]} + b^{[2]}$$

$$a^{[2]} = \sigma(z^{[2]})$$

$$W^{[2]} = \overset{[2]}{\underset{(1,4)}{\begin{bmatrix} w_1 & w_2 & w_3 & w_4 \end{bmatrix}}} \times \overset{[1]}{\underset{(4,1)}{\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}}} + \overset{[2]}{\underset{[1,1]}{b_1}}$$

Vetorização múltiplos exemplos

```
for i = 1 to m:
```

$$z^{[1](i)} = W^{[1]}\underline{x^{(i)}} + b^{[1]}$$

$$a^{[1](i)} = \sigma(z^{[1](i)})$$

$$z^{[2](i)} = W^{[2]}a^{[1](i)} + b^{[2]}$$

$$a^{[2](i)} = \sigma(z^{[2](i)})$$

$$\underline{Z^{[1]} = W^{[1]} \cdot X + b^{[1]}}$$

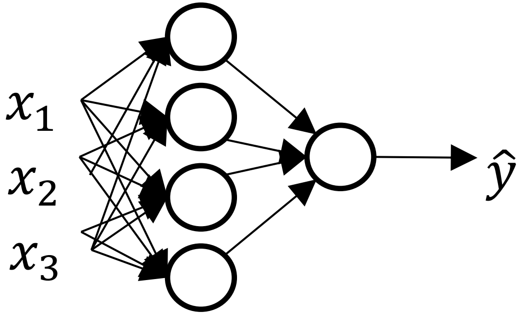
$$Z^{[1]} = \begin{bmatrix} | & | & & | \\ z^{[1](1)} & z^{[1](2)} & \cdots & z^{[1](m)} \\ | & | & & | \end{bmatrix}$$

$$Z^{[1]} = \underbrace{\underset{(4,3)}{W^{[1]}} \cdot \underset{(3,2)}{X}} + b^{[1]}$$

$$Z^{[1]} = \underset{(4,2)}{\begin{bmatrix} | & | \\ z^{[1](1)} & z^{[1](2)} \\ | & | \end{bmatrix}} + \underset{[4,1]}{b^{[1]}}$$

$$\underset{(4,2)}{\begin{bmatrix} | & | \\ b^{[1]} & b^{[1]} \\ | & | \end{bmatrix}}$$

Vetorização múltiplos exemplos



$$X = \begin{bmatrix} | & | & & | \\ x^{(1)} & x^{(2)} & \cdots & x^{(m)} \\ | & | & & | \end{bmatrix}$$

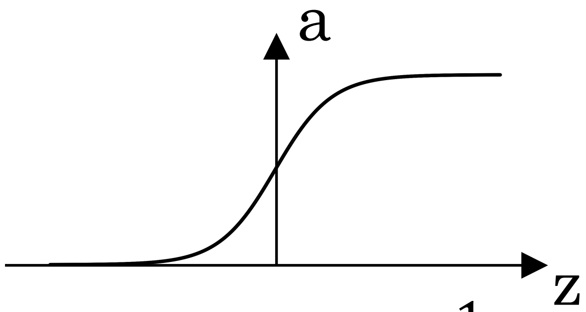
$$\mathrm{A}^{[1]} = \begin{bmatrix} | & | & & | \\ a^{[1](1)} & a^{[1](2)} & \cdots & a^{[1](m)} \\ | & | & & | \end{bmatrix}$$

```
for i = 1 to m
```

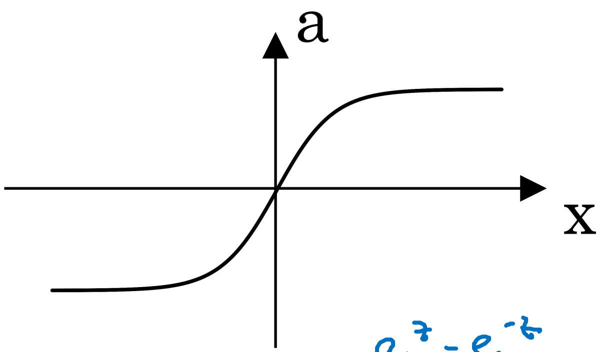
$$z^{[1](i)} = W^{[1]}x^{(i)} + b^{[1]}$$
$$a^{[1](i)} = \sigma(z^{[1](i)})$$
$$z^{[2](i)} = W^{[2]}a^{[1](i)} + b^{[2]}$$
$$a^{[2](i)} = \sigma(z^{[2](i)})$$

$$Z^{[1]} = W^{[1]}X + b^{[1]}$$
$$A^{[1]} = \sigma(Z^{[1]})$$
$$Z^{[2]} = W^{[2]}A^{[1]} + b^{[2]}$$
$$A^{[2]} = \sigma(Z^{[2]})$$

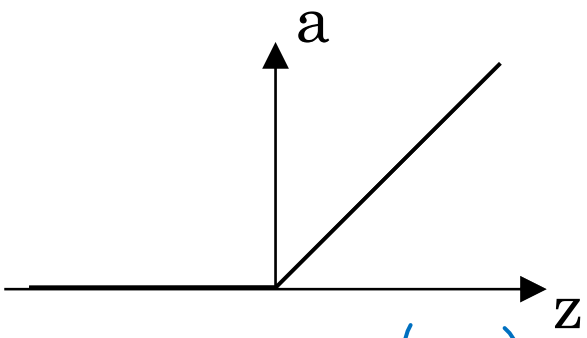
Pros e cons de funções de ativação



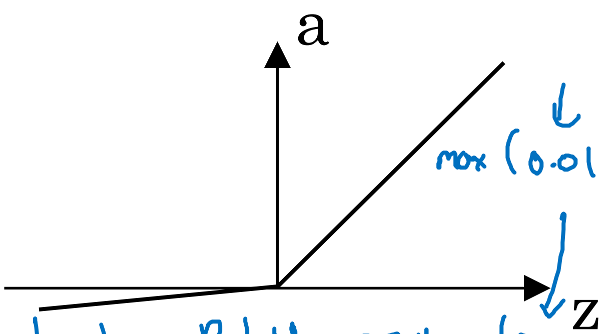
sigmoid: $a = \dfrac{1}{1+e^{-z}}$



tanh: $a = \dfrac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$



ReLU $a = \max(0, z)$

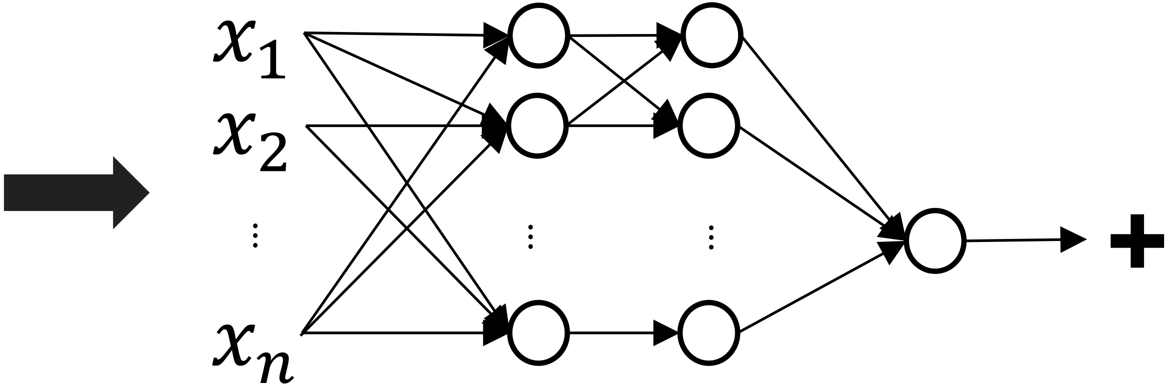


$\max(0.01z, z)$

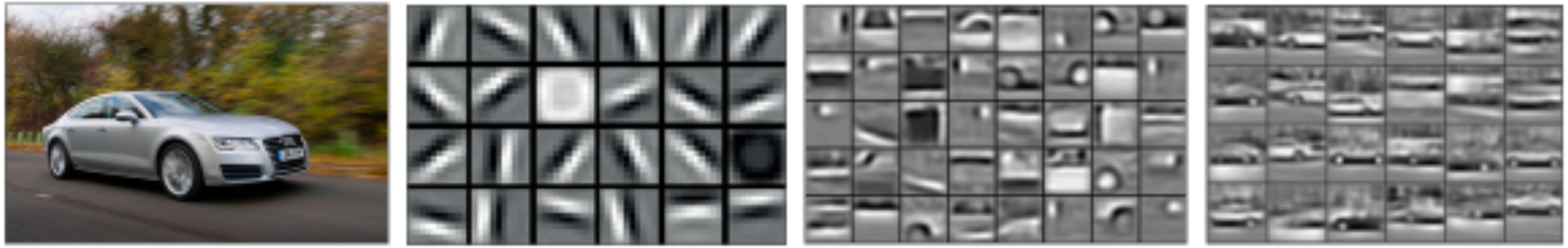
Leaky ReLU $a = \max(0.01z, z)$

Intuição

A comida
estava ótima.



Intuição



Referências

- Especialização em Machine Learning da Universidade de Washington:https://www.coursera.org/specializations/machine-learning
- Especialização em Deep Learning do Andrew Ng: https://www.coursera.org/specializations/deep-learning?