AULA 2
ESTRUTURA DAS REDES NEURAIS ARTIFICIAIS

1. Objetivos

- Mostrar analogia do sistema nervoso biológico com as RNAs.
- Apresentar estrutura das redes neurais artificiais (RNA):
  - Estrutura de um neurônio;
  - RNA de uma camada;
  - RNA de múltiplas camadas (deep-learning).
- Apresentar as funções de ativação.

2. Redes neurais biológicas

- Um neurônio biológico possui 4 componentes principais (Figura 1):
  - Dendritos;
  - Corpo celular;
  - Axônio;
  - Terminações dos axônios (sinapses).
- Dendritos recebem sinais provenientes de outros neurônios.
- O corpo celular soma os sinais de entrada $\Rightarrow$ quando a soma ultrapassa um determinado limite a célula dispara transmitindo um sinal para outros neurônios por meio do axônio.

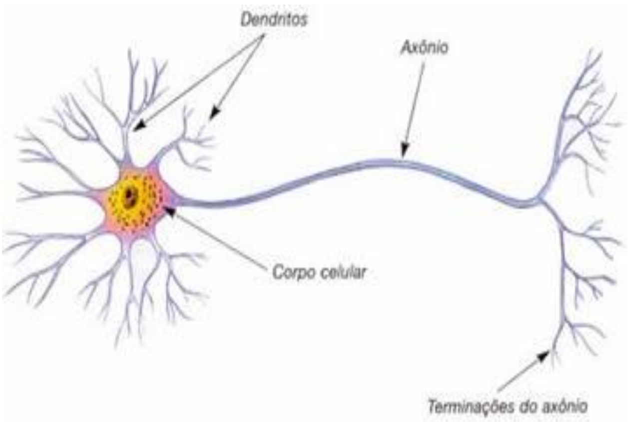


Figura1. Esquema de um neurônio biológico (Gomes, I., Visão Geral de RN - https://infosimples.com/artigos/visao-geral-redes-neurais).

3. Estrutura das RNAs

- Principais características das redes neurais biológicas estão presentes nas RNAs:
  - Uma RNA consiste de uma grande quantidade de unidades de processamento simples (neurônios) interconectados;
  - Cada neurônio artificial recebe muitos sinais;
  - Cada neurônio é conectado a outros neurônios por meio de ligações;
  - Os sinais recebidos pelos neurônios são modificados por um peso nas sinapses receptoras;
  - Os neurônios artificiais somam as entradas multiplicadas por ganhos (pesos);
  - O neurônio artificial transmite uma única saída;
  - A saída de um neurônio é transmitida para muitos outros neurônios;
  - Cada ligação entre neurônios possui o seu peso (ganho);
  - Uma RNA pode possuir várias camadas de neurônios.
- Tipos de camadas de uma RNA:
  - Entrada;
  - Intermediaria ou escondida;
  - Saída.
- Nas Figuras 2, 3 e 4 são mostrados exemplos de RNAs.

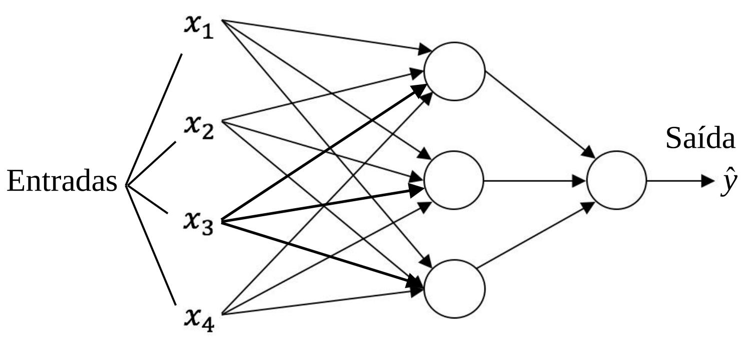


Figura 2. Exemplo de uma RNA de uma camada intermediária (escondida).

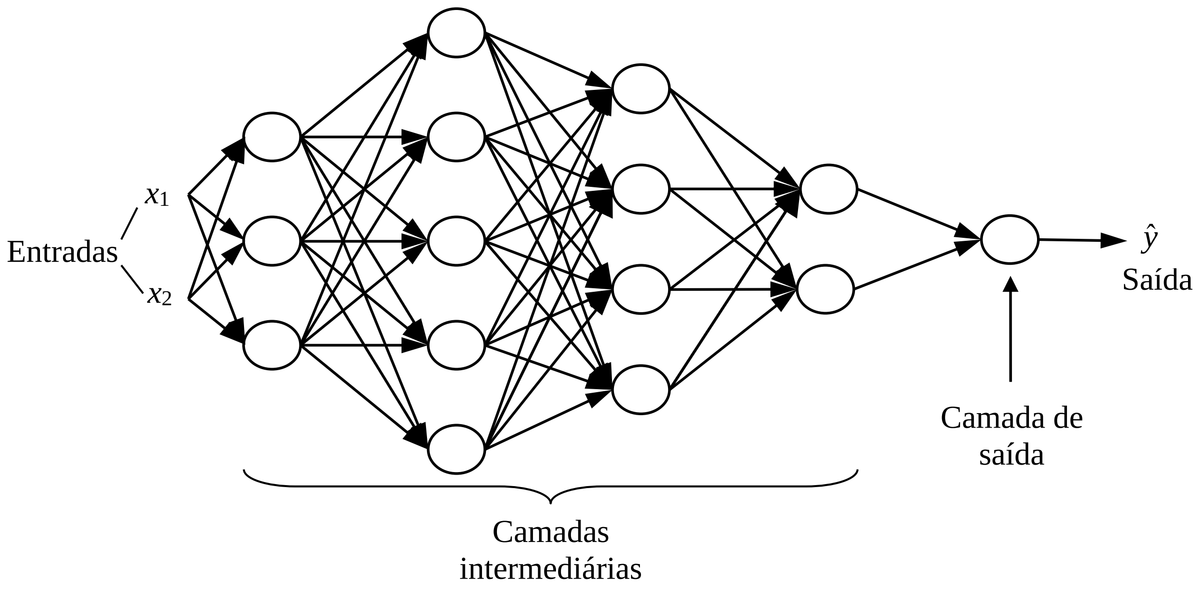


Figura 3. Exemplo de uma RNA de quatro camadas intermediárias (escondidas).

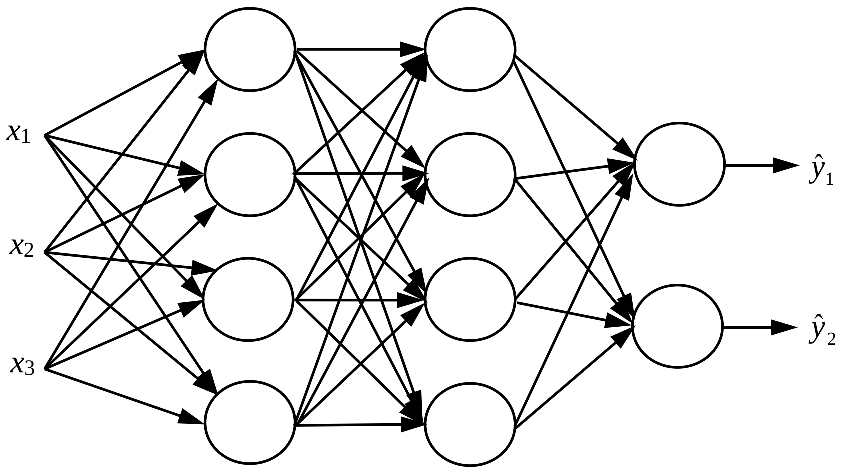


Figura 4. Exemplo de uma RNA de duas camadas intermediárias com duas saídas.

➢ Aprendizado de uma RNA:

- Os pesos das ligações representam a informação usada para resolver um problema;
- Aprender a solução de um problema ⇒ significa alterar os pesos das ligações;
- Aprendizado consiste em modificar os pesos das ligações de forma iterativa até a RNA ser capaz de reproduzir os dados usados no treinamento.

4. Estrutura de um neurônio

- Na Figura 5 é apresentado um esquema de um neurônio da primeira camada de uma RNA.
- Cada neurônio tem seu estado interno (z) $\Rightarrow$ que é função das entradas recebidas.
- Cada neurônio tem o seu nível de ativação (a) $\Rightarrow$ que é função das entradas recebidas, dos pesos de suas ligações e da sua função de ativação.
- Força da conexão entre neurônios consiste de um ganho específico para cada conexão.

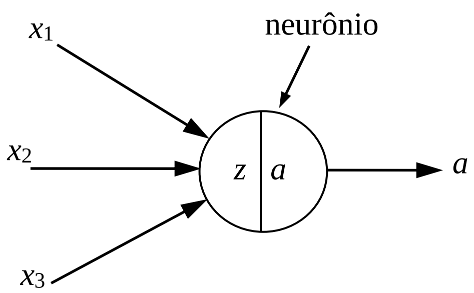


Figura 5. Esquema de um neurônio da primeira camada intermediária.

- **Equações do neurônio da primeira camada intermediária:**
  - Estado do neurônio (z) $\Rightarrow$ somatória ponderada das entradas (no caso são as entradas da RNA):

$$\boxed{z = w_1 x_1 + w_2 x_2 + w_3 x_3 + b} \tag{1}$$

onde:

w_j = peso da ligação do neurônio com a entrada j

b = viés do neurônio

  - Nível de ativação do neurônio (a):

$$\boxed{a = g(z)} \tag{2}$$

onde $g(z)$ = função de ativação do neurônio

- **Matricialmente:**
  - Dimensões:
    - Número de entradas do neurônio $\Rightarrow$ n_x (no caso é igual ao número de entradas da RNA)

- Vetor de entradas $\Rightarrow$ dimensão (n_x, 1)
- Vetor de pesos $\Rightarrow$ dimensão (1, n_x)
- Viés, estado, nível de ativação e saída do neurônio $\Rightarrow$ escalares

- Equações:

$$\boxed{\begin{cases} z = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + b \Rightarrow z = \mathbf{wx} + b \\ \\ a = g(z) \end{cases}} \tag{3}$$

onde:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{(n_x,1)} \quad \Rightarrow \text{vetor de entradas}$$

$$\mathbf{w} = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}_{(1,n_x)} \quad \Rightarrow \text{vetor de pesos das sinapses}$$

Cada neurônio possui um vetor linha de pesos, com dimensão igual ao número de entradas do neurônio.

4. RNA de uma única camada intermediária (RNA rasa)

➢ Na Figura 6 é apresentada uma RNA de uma única camada intermediária (RNA rasa).

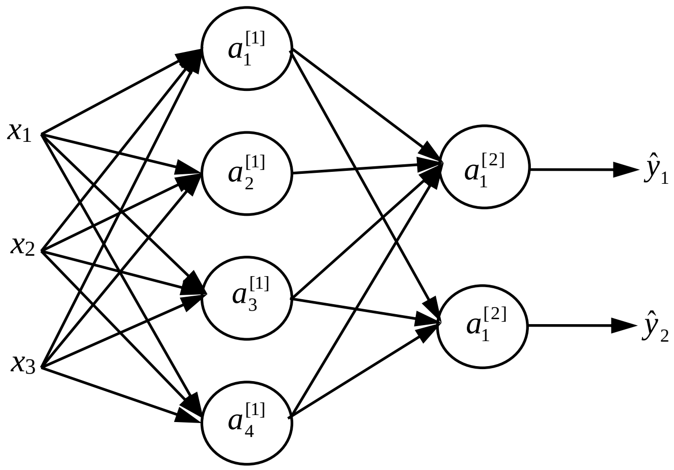


Figura 6. RNA rasa de uma camada intermediária com 4 neurônios, 3 entradas e 2 saídas.

- **Equações da RNA:**

- Dimensões:
  - Número de entradas ⇒ $n_x = 3$
  - Número de neurônios das camadas:
    - Primeira camada ⇒ $n^{[1]} = 4$
    - Segunda camada ⇒ $n^{[2]} = 2$
    - Número de estados = número de níveis de ativação = número de vieses = número de neurônios da camada
  - Matriz de pesos:
    - Primeira camada ⇒ dimensão $(n^{[1]}, n_x) = (4, 3)$
    - Segunda camada ⇒ dimensão $(n^{[2]}, n^{[1]}) = (2, 4)$
  - Número de saídas ⇒ n_y = número de neurônios da 2ª camada ($n^{[2]} = 2$)

- Estados dos neurônios da primeira camada ($\mathbf{z}^{[1]}$):

$$\begin{cases} z_1^{[1]} = w_{1,1}^{[1]}x_1 + w_{1,2}^{[1]}x_2 + w_{1,3}^{[1]}x_3 + b_1^{[1]} \\ z_2^{[1]} = w_{2,1}^{[1]}x_1 + w_{2,2}^{[1]}x_2 + w_{2,3}^{[1]}x_3 + b_2^{[1]} \\ z_3^{[1]} = w_{3,1}^{[1]}x_1 + w_{3,2}^{[1]}x_2 + w_{3,3}^{[1]}x_3 + b_3^{[1]} \\ z_4^{[1]} = w_{4,1}^{[1]}x_1 + w_{4,2}^{[1]}x_2 + w_{4,3}^{[1]}x_3 + b_4^{[1]} \end{cases} \Rightarrow \begin{bmatrix} z_1^{[1]} \\ z_2^{[1]} \\ z_3^{[1]} \\ z_4^{[1]} \end{bmatrix} = \begin{bmatrix} w_{1,1}^{[1]} & w_{1,2}^{[1]} & w_{1,3}^{[1]} \\ w_{2,1}^{[1]} & w_{2,2}^{[1]} & w_{2,3}^{[1]} \\ w_{3,1}^{[1]} & w_{3,2}^{[1]} & w_{3,3}^{[1]} \\ w_{4,1}^{[1]} & w_{4,2}^{[1]} & w_{4,3}^{[1]} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b_1^{[1]} \\ b_2^{[1]} \\ b_3^{[1]} \\ b_4^{[1]} \end{bmatrix} \tag{4}$$

ou simplesmente:

$$\boxed{\mathbf{z}^{[1]} = \mathbf{W}^{[1]}\mathbf{x} + \mathbf{b}^{[1]}} \tag{5}$$

onde:

$w_{k,j}^{[1]}$ = peso da ligação do neurônio k da primeira camada com a entrada j

$b_k^{[1]}$ = viés do neurônio k da primeira camada

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{(n_x,1)}$$ = vetor de entradas

$$\mathbf{z}^{[1]} = \begin{bmatrix} z_1^{[1]} \\ z_2^{[1]} \\ z_3^{[1]} \\ z_4^{[1]} \end{bmatrix}_{(n^{[1]},1)}$$ = vetor de estados dos neurônios da primeira camada

$$\mathbf{W}^{[1]} = \begin{bmatrix} w_{1,1}^{[1]} & w_{1,2}^{[1]} & w_{1,3}^{[1]} \\ w_{2,1}^{[1]} & w_{2,2}^{[1]} & w_{2,3}^{[1]} \\ w_{3,1}^{[1]} & w_{3,2}^{[1]} & w_{3,3}^{[1]} \\ w_{4,1}^{[1]} & w_{4,2}^{[1]} & w_{4,3}^{[1]} \end{bmatrix}_{(n^{[1]},n_x)}$$ = matriz de pesos da primeira camada

$$\mathbf{b}^{[1]} = \begin{bmatrix} b_1^{[1]} \\ b_2^{[1]} \\ b_3^{[1]} \\ b_4^{[1]} \end{bmatrix}_{(n^{[1]},1)}$$ = vetor de vieses dos neurônios da primeira camada

- Nível de ativação dos neurônios da primeira camada ($\mathbf{a}^{[1]}$):

$$\begin{cases} a_1^{[1]} = g^{[1]}(z_1^{[1]}) \\ a_2^{[1]} = g^{[1]}(z_2^{[1]}) \\ a_3^{[1]} = g^{[1]}(z_3^{[1]}) \\ a_4^{[1]} = g^{[1]}(z_4^{[1]}) \end{cases} \Rightarrow \begin{bmatrix} a_1^{[1]} \\ a_2^{[1]} \\ a_3^{[1]} \\ a_4^{[1]} \end{bmatrix} = g^{[1]}\left(\begin{bmatrix} z_1^{[1]} \\ z_2^{[1]} \\ z_3^{[1]} \\ z_4^{[1]} \end{bmatrix} \right) \Rightarrow \boxed{\mathbf{a}^{[1]} = g^{[1]}(\mathbf{z}^{[1]})} \tag{6}$$

onde:

$g^{[1]}$ = função de ativação da primeira camada

$$\mathbf{a}^{[1]} = \begin{bmatrix} a_1^{[1]} \\ a_2^{[1]} \\ a_3^{[1]} \\ a_4^{[1]} \end{bmatrix}_{(n^{[1]},1)}$$ = vetor de ativação dos neurônios da primeira camada

- Estados dos neurônios da segunda camada ($\mathbf{z}^{[2]}$):

$$\begin{cases} z_1^{[2]} = w_{1,1}^{[2]} a_1^{[1]} + w_{1,2}^{[2]} a_2^{[1]} + w_{1,3}^{[2]} a_3^{[1]} + w_{1,4}^{[2]} a_4^{[1]} + b_1^{[2]} \\ z_2^{[2]} = w_{2,1}^{[2]} a_1^{[1]} + w_{2,2}^{[2]} a_2^{[1]} + w_{2,3}^{[2]} a_3^{[1]} + w_{2,4}^{[2]} a_4^{[1]} + b_2^{[2]} \end{cases} \tag{7}$$

Matricialmente,

$$\begin{bmatrix} z_1^{[2]} \\ z_2^{[2]} \end{bmatrix} = \begin{bmatrix} w_{1,1}^{[2]} & w_{1,2}^{[2]} & w_{1,3}^{[2]} & w_{1,4}^{[2]} \\ w_{2,1}^{[2]} & w_{2,2}^{[2]} & w_{2,3}^{[2]} & w_{2,4}^{[2]} \end{bmatrix} \begin{bmatrix} a_1^{[1]} \\ a_2^{[1]} \\ a_3^{[1]} \\ a_4^{[1]} \end{bmatrix} + \begin{bmatrix} b_1^{[2]} \\ b_2^{[2]} \end{bmatrix} \Rightarrow \boxed{\mathbf{z}^{[2]} = \mathbf{W}^{[2]}\mathbf{a}^{[1]} + \mathbf{b}^{[2]}} \tag{8}$$

onde:

$w_{k,j}^{[2]}$ = peso da ligação do neurônio *k* da camada 2 com o neurônio *j* da primeira camada

$b_k^{[2]}$ = viés do neurônio k da segunda camada

$$\mathbf{z}^{[2]} = \begin{bmatrix} z_1^{[2]} \\ z_2^{[2]} \end{bmatrix}_{(n^{[2]},1)}$$ = vetor de estados dos neurônios da segunda camada

$$\mathbf{W}^{[2]} = \begin{bmatrix} w_{1,1}^{[2]} & w_{1,2}^{[2]} & w_{1,3}^{[2]} & w_{1,4}^{[2]} \\ w_{2,1}^{[2]} & w_{2,2}^{[2]} & w_{2,3}^{[2]} & w_{2,4}^{[2]} \end{bmatrix}_{(n^{[2]},n^{[1]})}$$ = matriz de pesos da segunda camada

$$\mathbf{b}^{[2]} = \begin{bmatrix} b_1^{[2]} \\ b_2^{[2]} \end{bmatrix}_{(n^{[2]},1)}$$ = vetor de vieses dos neurônios da segunda camada

- Nível de ativação dos neurônios da segunda camada ($\mathbf{a}^{[2]}$):

$$\begin{cases} a_1^{[2]} = g^{[2]}(z_1^{[2]}) \\ a_2^{[2]} = g^{[2]}(z_2^{[2]}) \end{cases} \Rightarrow \begin{bmatrix} a_1^{[2]} \\ a_2^{[2]} \end{bmatrix} = g^{[2]}\left(\begin{bmatrix} z_1^{[2]} \\ z_2^{[2]} \end{bmatrix}\right) \Rightarrow \boxed{\mathbf{a}^{[2]} = g^{[2]}(\mathbf{z}^{[2]})} \tag{9}$$

onde:

$g^{[2]}$ = função de ativação da segunda camada

$$\mathbf{a}^{[2]} = \begin{bmatrix} a_1^{[2]} \\ a_2^{[2]} \end{bmatrix}_{(n^{[2]},1)}$$ = vetor de ativação dos neurônios da segunda camada

$\Rightarrow$ Observe que as funções de ativações das diversas camadas em geral são diferentes.

- Saídas da rede ($\hat{\mathbf{y}}$):

$$\begin{cases} \hat{y}_1 = a_1^{[2]} \\ \hat{y}_2 = a_2^{[2]} \end{cases} \Rightarrow \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = \begin{bmatrix} a_1^{[2]} \\ a_2^{[2]} \end{bmatrix} \Rightarrow \boxed{\hat{\mathbf{y}} = \mathbf{a}^{[2]}} \tag{10}$$

onde:

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix}_{(n^{[2]},1)}$$ = vetor de saída da RN

➢ Na Figura 7 é apresentado o fluxo de dados em uma RNA rasa.

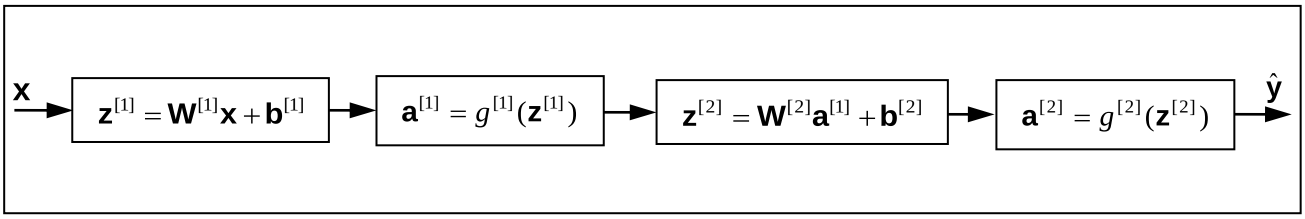


Figura 7. Fluxo de cálculo da RNA rasa de uma camada intermediária.

5. <u>RN de múltiplas camadas intermediárias (Deep-learning)</u>

➢ Na Figura 8 é apresentada uma RNA deep-learning (ou profunda).

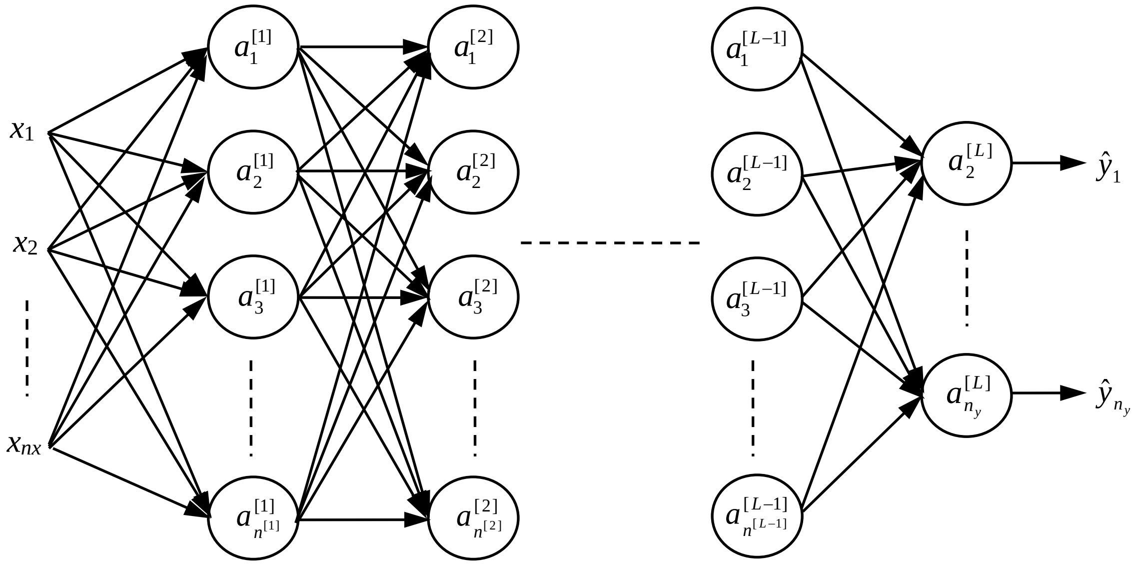


Figura 8. RNA deep-learning de *L* camadas.

➢ **Parâmetros e dimensões:**

- Número de entradas ⇒ n_x
- Vetor de entradas ⇒ $\mathbf{x}$, dimensão (n_x, 1)
- Número de saídas ⇒ n_y
- Vetor de saídas ⇒ $\hat{\mathbf{y}}$, dimensão (n_y, 1)
- Número total de camadas ⇒ L (número de camadas intermediárias ⇒ $L - 1$)
- Número de neurônios da l-ésima camada ⇒ $n^{[l]}$
- Vetor de estados dos neurônios da l-ésima camada ⇒ $\mathbf{z}^{[l]}$, dimensão ($n^{[l]}$, 1)
- Vetor de ativações dos neurônios da l-ésima camada ⇒ $\mathbf{a}^{[l]}$, dimensão ($n^{[l]}$, 1)

- Matriz de pesos da *l*-ésima camada ⇒ $\mathbf{W}^{[l]}$, dimensão ($n^{[l]}$, $n^{[l-1]}$)
- Vetor de vieses da *l*-ésima camada ⇒ $\mathbf{b}^{[l]}$, dimensão ($n^{[l]}$, 1)

➢ **Equações**

- Primeira camada:

$$\mathbf{z}^{[1]} = \mathbf{W}^{[1]}\mathbf{x} + \mathbf{b}^{[1]} \tag{11}$$

$$\mathbf{a}^{[1]} = g^{[1]}(\mathbf{z}^{[1]}) \tag{12}$$

- *l*-ésima camada (l = 2,..., L):

$$\mathbf{z}^{[l]} = \mathbf{W}^{[l]}\mathbf{a}^{[l-1]} + \mathbf{b}^{[l]} \tag{13}$$

$$\mathbf{a}^{[l]} = g^{[l]}(\mathbf{z}^{[l]}) \tag{14}$$

- Saídas:

$$\hat{\mathbf{y}} = \mathbf{a}^{[L]} \tag{15}$$

➢ O fluxo de dados em uma RNA deep-learning de *L* camadas é apresentado na Figura 9.

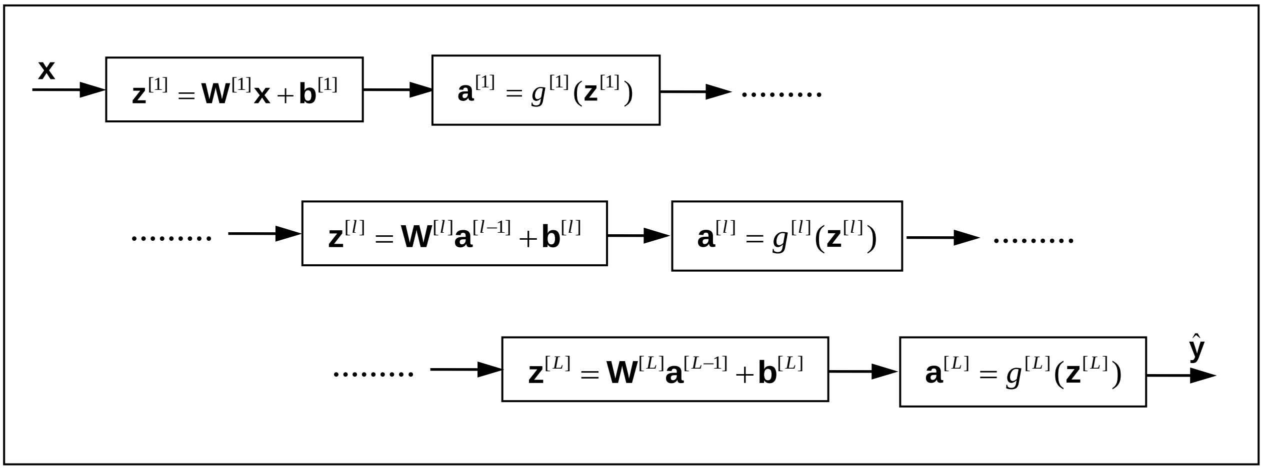


Figura 9. Fluxo de cálculo de uma RNA deep-learning de *L* camadas.

6. <u>Nomenclatura geral</u>

➢ **Para facilitar o entendimento usamos uma notação matemática padrão para as RNAs.**

➢ Comentários gerais:

- Superescrito (i) $\Rightarrow$ i-ésimo exemplo de treinamento
- Superescrito [l] $\Rightarrow$ l-ésima camada
- Subescrito k $\Rightarrow$ k-ésimo elemento de um vetor
- Subescritos k, j $\Rightarrow$ índice do elemento de uma matriz na k-ésima linha e j-ésima coluna

➢ Dimensões:

- $n_x \Rightarrow$ número de entradas de cada exemplo
- $m \Rightarrow$ número total de exemplos usados no treinamento
- $n^{[l]} \Rightarrow$ número de neurônios da l-ésima camada
- $L \Rightarrow$ número de camadas da RNA
- $n_y = n^{[L]} \Rightarrow$ número de saídas da RNA = número de neurônios da última camada (L)

➢ Parâmetros:

- $\mathbf{x}^{(i)} \Rightarrow$ i-ésimo vetor de entrada $\rightarrow$ dimensão (n_x,1)
- $\mathbf{X} \Rightarrow$ matriz com todos os vetores de entradas $\rightarrow$ dimensão (n_x, m)
- $\mathbf{y}^{(i)} \Rightarrow$ i-ésimo vetor de saída $\rightarrow$ dimensão (n_y, 1)
- $\mathbf{Y} \Rightarrow$ matriz com todos os vetores de saídas $\rightarrow$ dimensão (n_y, m)
- $\mathbf{z}^{[l](i)} \Rightarrow$ vetor de estados dos neurônios da l-ésima camada referentes ao i-ésimo vetor de entrada $\rightarrow$ dimensão ($n^{[l]}$, 1)
- $\mathbf{Z}^{[l]} \Rightarrow$ matriz com todos os estados da l-ésima camada para todas as entradas $\rightarrow$ dimensão ($n^{[l]}$, m)
- $\mathbf{a}^{[l](i)} \Rightarrow$ vetor de ativações dos neurônios da l-ésima camada referentes ao i-ésimo vetor de entrada $\rightarrow$ dimensão ($n^{[l]}$, 1)
- $\mathbf{A}^{[l]} \Rightarrow$ matriz com todas as ativações da l-ésima camada para todas as entradas $\rightarrow$ dimensão ($n^{[l]}$, m)
- $\mathbf{W}^{[l]} \Rightarrow$ matriz de pesos da l-ésima camada $\rightarrow$ dimensão ($n^{[l]}$, $n^{[l-1]}$)
- $\mathbf{b}^{[l]} \Rightarrow$ vetor de vieses da l-ésima camada $\rightarrow$ dimensão ($n^{[l]}$, 1)

7. <u>Funções de ativação</u>

➢ Existem diversos tipos de funções de ativação.

➢ Funções de ativação mais usadas:

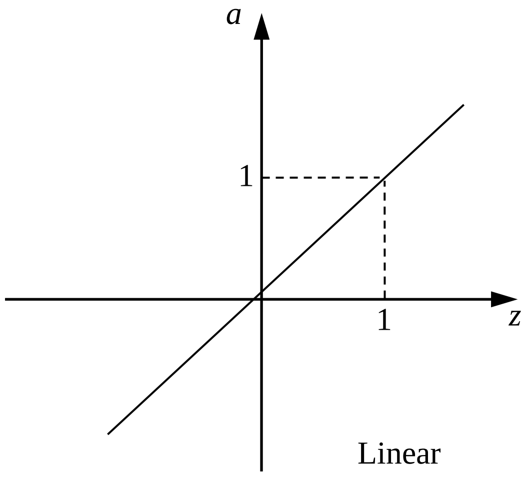
- Linear;
- Sigmóide;
- Tangente hiperbólica;

- Unidade de retificação linear ("Retified Linear Unit" – **ReLu**);
- Unidade de retificação linear fraca ("Leaky Retified Linear Unit" – **LeReLu**).

➢ Todos os neurônios de uma camada têm que ter a mesma função de ativação, mas cada camada da RNA pode ter uma função de ativação diferente.

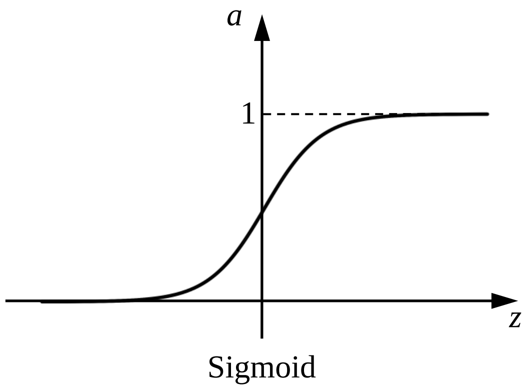
➢ O tipo de função de ativação usado em uma RNA depende do problema que se deseja resolver.

Função linear



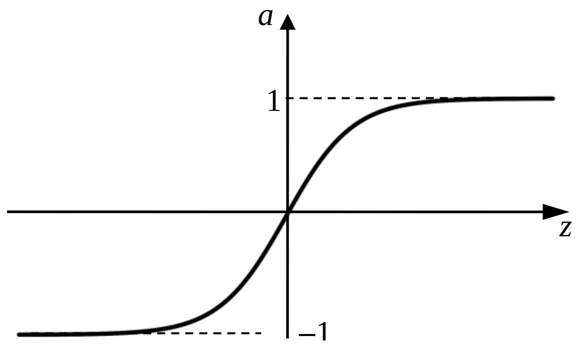
$$\boxed{a = z} \tag{16}$$

Função sigmóide



$$\boxed{a = \sigma(z) = \frac{1}{1 + e^{-z}}} \tag{17}$$

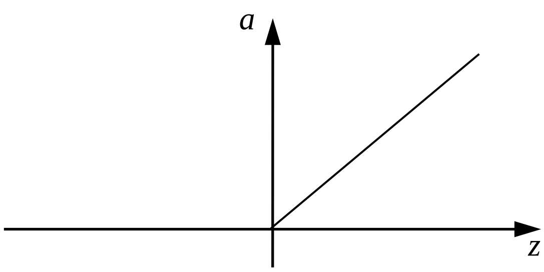
Função tangente hiperbólica



Tangente hiperbólica

$$a = \tanh(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \tag{18}$$

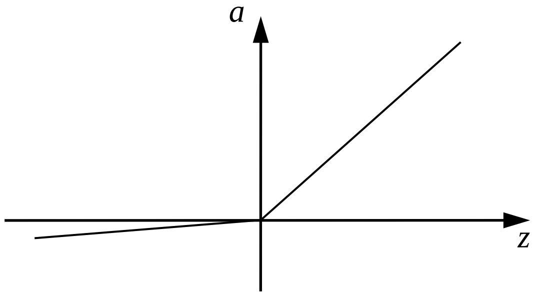
Função ReLu



ReLu

$$a = \begin{cases} 0, & \text{para } z < 0 \\ z, & \text{para } z \geq 0 \end{cases} \qquad a = \max(0, z) \tag{19}$$

Função LeReLu



LeReLu

$$a = \begin{cases} 0{,}01z, & \text{para } z < 0 \\ z, & \text{para } z \geq 0 \end{cases} \qquad a = \max(0{,}01z, z) \tag{20}$$

- **Uma RNA não pode ter todas as suas camadas com funções de ativação linear.**
- **O que aconteceria se todas as funções de ativação fossem lineares?**

  Para responder essa pergunta vamos analisar a RN rasa mostrada na Figura 10.

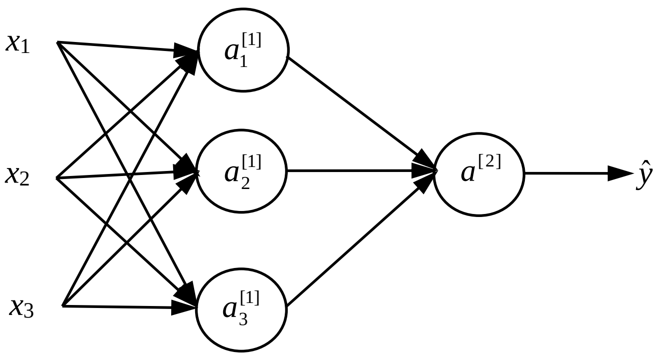


Figura 10. Esquema de uma RNA rasa com 3 neurônios na camada intermediária.

Dado o vetor de entradas $\mathbf{x}$, a saída da RNA da Figura 10, $\hat{y}$, é dada por:

$$\begin{cases} \mathbf{z}^{[1]} = \mathbf{W}^{[1]}\mathbf{x} + \mathbf{b}^{[1]} \\ \mathbf{a}^{[1]} = g^{[1]}(\mathbf{z}^{[1]}) \\ z^{[2]} = \mathbf{W}^{[2]}\mathbf{a}^{[1]} + b^{[2]} \\ \hat{y} = a^{[2]} = g^{[2]}(z^{[2]}) \end{cases} \tag{21}$$

Se todas as funções de ativação fossem funções lineares, ou seja:

$$g^{[1]}(z) = g^{[2]}(z) = z \tag{22}$$

Tem-se para a primeira camada:

$$\mathbf{a}^{[1]} = \mathbf{z}^{[1]} = \mathbf{W}^{[1]}\mathbf{x} + \mathbf{b}^{[1]} \tag{23}$$

Para a segunda camada, tem-se:

$$a^{[2]} = z^{[2]} = \mathbf{W}^{[2]}\mathbf{a}^{[1]} + b^{[2]} \tag{24}$$

Substituindo $\mathbf{a}^{[1]}$ na equação (24), obtém-se:

$$a^{[2]} = \mathbf{W}^{[2]}\left[\mathbf{W}^{[1]}\mathbf{x} + \mathbf{b}^{[1]}\right] + b^{[2]} \tag{25}$$

ou seja:

$$a^{[2]} = \underbrace{\mathbf{W}^{[2]}\mathbf{W}^{[1]}\mathbf{x}}_{\mathbf{W}'\mathbf{x}} + \underbrace{\mathbf{W}^{[2]}\mathbf{b}^{[1]} + b^{[2]}}_{b'} \tag{26}$$

Portanto, a saída de RNA é uma função linear dada por:

$$\hat{y} = a^{[2]} = \mathbf{W}'\mathbf{x} + b' \Rightarrow \textbf{função linear} \tag{27}$$

⇒ **Conclusão:**

- **Se as funções de ativação forem todas lineares a saída será sempre uma combinação linear das entradas**

- **RNA somente aprenderia funções lineares**

- **RNA muito limitada**

8. Aprendizado (treinamento) das RNAs

➢ Para uma RNA ser capaz de resolver algum problema ela precisa ser treinada.

➢ O treinamento ou aprendizado consiste em determinar os parâmetros da RNA (matrizes $\mathbf{W}^{[l]}$ e vetores $\mathbf{b}^{[l]}$, para $l = 1,...,L$) para que a RNA seja capaz de resolver o problema desejado.

➢ **A pergunta que surge é: Como determinar os pesos da RNA?**

Por meio de um processo de otimização onde iterativamente são calculados os parâmetros para minimizar uma determinada função de custo.