

Redes Neurais Artificiais – Parte I

1. Prólogo

- A descoberta da **célula** (Robert Hooke, 1665) foi um passo de enorme importância para que houvesse uma melhor compreensão da estrutura dos seres vivos. Talvez se possa considerar, *cum grano salis*, a célula um “átomo de vida”.
- As células eucariontes possuem três partes principais: membrana, núcleo e citoplasma. A membrana “delimita a célula”, *i.e.*, isola seu interior do meio externo. O núcleo abriga o material genético e, no citoplasma, estão componentes como as organelas.

- Os neurônios são células, mas células que possuem mecanismos elétricos e/ou químicos peculiares. A Fig. 1 traz uma visão esquemática.

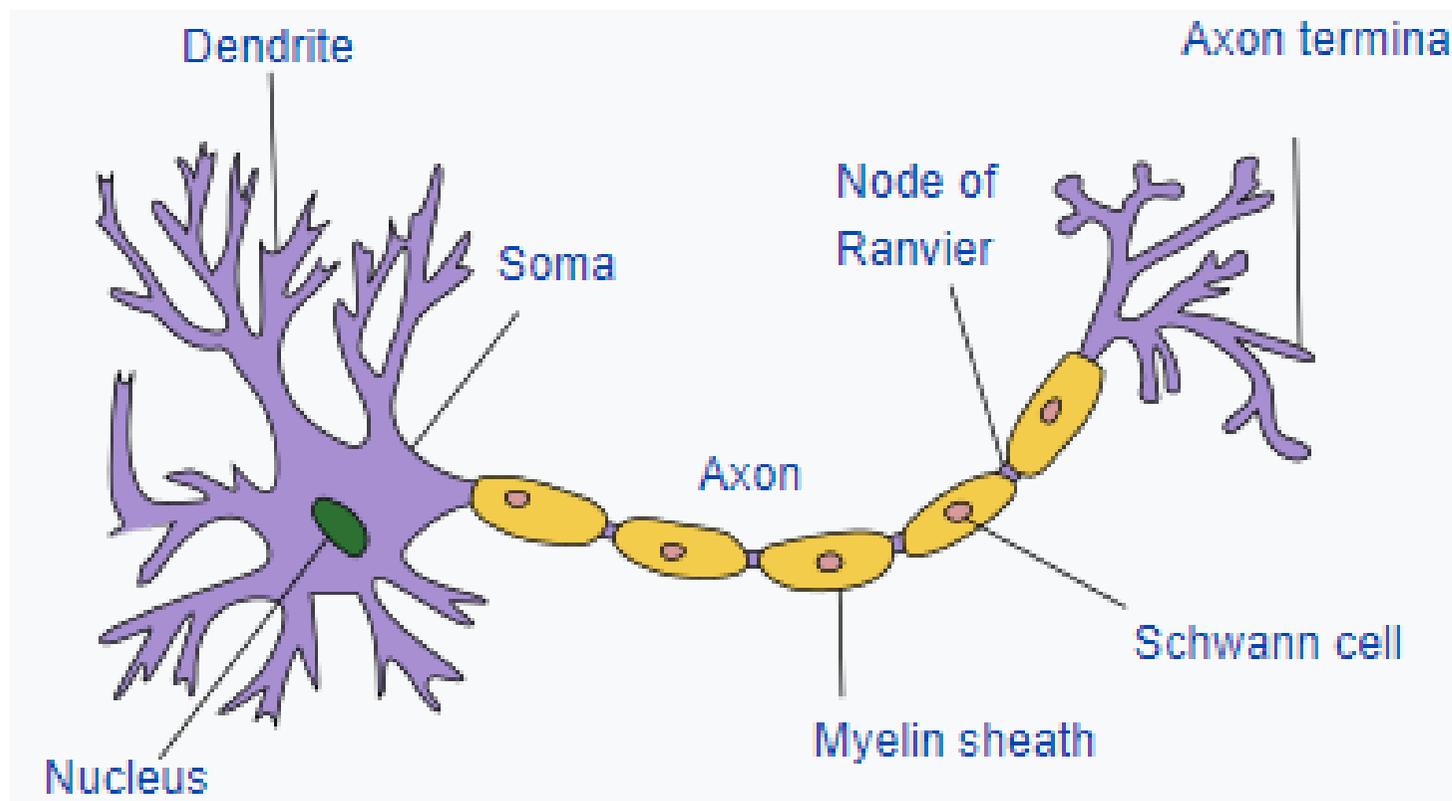


Figura 1 – Visão Básica de um Neurônio.

- Em termos muito simples, podemos afirmar (cientes de que há exceções):

- Que o neurônio recebe estímulos elétricos, basicamente a partir dos dendritos;
- Que esses estímulos são integrados;
- Que a estimulação pode levar à geração de uma resposta elétrica enviada pelo axônio.
- Do ponto de vista de nosso curso, o neurônio será um sistema com entradas e saída. Como é a regra quando se lida com sistemas, o neurônio **processa informação**.
- Uma “perspectiva de trabalho” pode ser a seguinte:
 - Os neurônios recebem estímulos elétricos;
 - Esses estímulos são integrados;
 - Se a atividade exceder certo limiar, o neurônio gera um pulso (*spike* ou potencial de ação);

O potencial de ação é mostrado na Fig. 2.

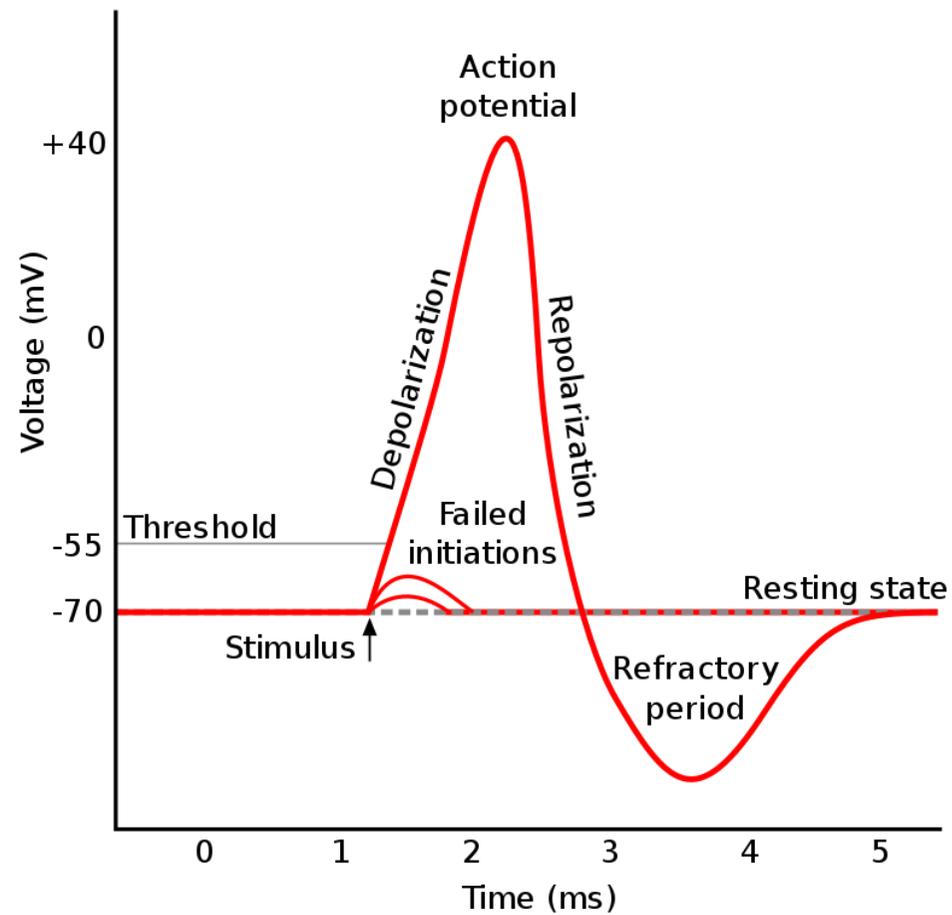


Figura 2 – Ilustração do Potencial de Ação.

2. O Modelo de McCulloch e Pitts

- O final do século XIX e o início do século XX foram períodos fundamentais para o estabelecimento da visão atual do sistema nervoso. Quando Warren McCulloch e Walter Pitts apresentaram o primeiro modelo “computacional” de neurônio (MCCULLOCH & PITTS, 1943), já haviam sido identificados vários elementos considerados centrais até hoje.
- O modelo de McCulloch e Pitts parece “simples” à luz de alguns modelos empregados nos dias atuais, mas sua importância foi enorme. A partir desse modelo, foi possível estabelecer uma conexão entre o funcionamento de um neurônio e a lógica proposicional. A partir daí, a relação com a computação digital foi natural.
- As premissas do modelo são (MCCULLOCH & PITTS, 1943):

- A atividade do neurônio é um processo do tipo “tudo ou nada”, ou seja, um processo binário.
- Certo número de sinapses deve ser excitado num determinado período para que o neurônio “dispare”.
- O único atraso significativo no sistema nervoso é o atraso sináptico.
- A atividade numa sinapse inibitória impede o disparo do neurônio no instante associado.
- A estrutura da rede de neurônios não se altera com o tempo.
- Um exemplo desse modelo está na Fig. 3. Note que o limiar de disparo é dois e que a terceira entrada se liga por meio de uma sinapse inibitória.

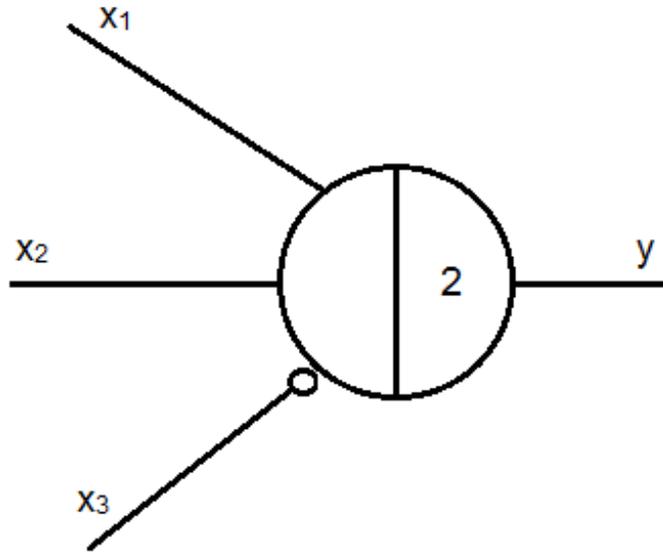


Figura 3 – Exemplo de Modelo de McCulloch e Pitts.

- A tabela-verdade que rege o funcionamento desse neurônio é mostrada a seguir.

Tabela 1 – Tabela-Verdade do Exemplo (Fig. 3).

x_1	x_2	x_3	y
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	1
1	1	1	0

3. Perceptron

- O modelo conhecido como *perceptron*, proposto por Frank Rosenblatt no final dos anos 1950, foi um dos grandes marcos na história das redes neurais. Sua estrutura geral é similar à do modelo de McCulloch e Pitts, mas há diferenças fundamentais, como o abandono de um domínio numérico exclusivamente booleano (binário). A Fig. 4 traz uma representação do modelo.

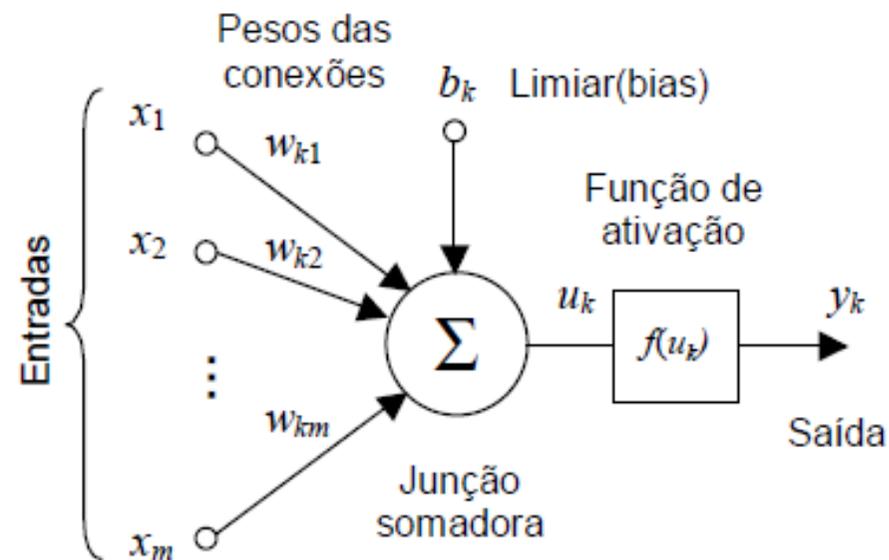


Figura 4 – Estrutura do perceptron.

- A ideia é que a ativação do neurônio (causada pelos estímulos de entrada) seja uma combinação linear entre os estímulos e os pesos sinápticos. Se essa ativação exceder certo limiar, ocorrerá um “disparo”. Isso pode ser expresso por meio de uma **função de ativação** do tipo “degrau”. Ou seja, de acordo com a notação da Fig. 4,

$$y_k = f(u_k) = f\left(\sum_{i=1}^m w_{ki}x_i + b_k\right) = \begin{cases} 1, & \text{se } \sum_{i=1}^m w_{ki}x_i + b_k \geq 0 \\ 0, & \text{se } \sum_{i=1}^m w_{ki}x_i + b_k < 0 \end{cases}$$

- Note que a função de ativação $f(\cdot)$ está “em torno de zero”, e o limiar de disparo é controlado, indiretamente, pelo valor do bias (b_k).

- O tipo de resposta desse neurônio dá origem a um classificador para problemas com duas classes. As classes são separadas pela fronteira de pontos para os quais vale:

$$\sum_{i=1}^m w_{ki}x_i + b_k = 0$$

- No espaço dos atributos x_i , essa é a equação de um hiperplano. Esse ponto é muito importante: o perceptron só é capaz de classificar dados que sejam linearmente separáveis (separáveis por um hiperplano). A Fig. 5 ilustra isso para um caso bidimensional.

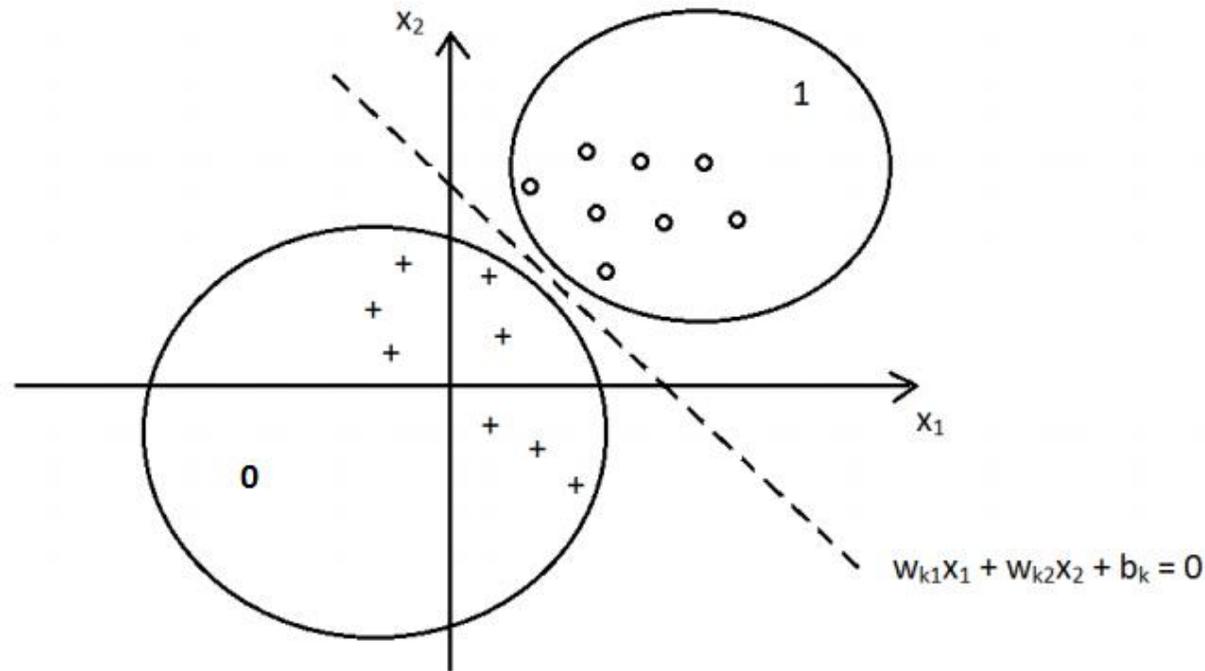


Figura 5 – Dados Linearmente Separáveis.

- Caso os dados sejam linearmente separáveis, a **regra de aprendizado do perceptron** tem convergência garantida num número finito de iterações. Nessa regra, para cada dado do conjunto de treino, obtém-se, primeiramente, a saída da rede para os pesos sinápticos atuais:

$$y = f\left(\sum_{i=1}^m w_i x_i + b\right) = f(\mathbf{w}^T \mathbf{x} + b)$$

- Em seguida, gera-se um indicador de erro entre a saída e o rótulo d do dado:

$$e = d - y$$

- Caso esse erro seja não-nulo, a adaptação é da seguinte forma:

$$\mathbf{w} \leftarrow \mathbf{w} + \alpha e \mathbf{x}$$

$$b \leftarrow b + \alpha e$$

sendo α uma taxa de aprendizado.

- Após a apresentação de todos os dados (ou seja, uma **época**), deve haver uma reordenação dos padrões e uma nova apresentação. Quando a separação linear ocorrer, não haverá mais erros, e as regras de atualização dadas não mais modificarão os parâmetros.

4. Referências bibliográficas

MCCULLOCH, W., PITTS, W., “A Logical Calculus of the Ideas Immanent in Nervous Activity”,
The Bulletin of Mathematical Biophysics, Vol. 5, No. 4, pp. 115 – 133, 1943.

VON ZUBEN, F. J., **Notas de Aulas do Curso “Redes Neurais” (IA353)**, disponíveis em
<http://www.dca.fee.unicamp.br/~vonzuben/courses/ia353.html>

WIKIPEDIA, **Artigos e Figuras Diversas**.