

Aprendizado de Máquina e Deep Learning

Redes Neurais Artificiais

Prof. Dr. Thiago Meirelles Ventura

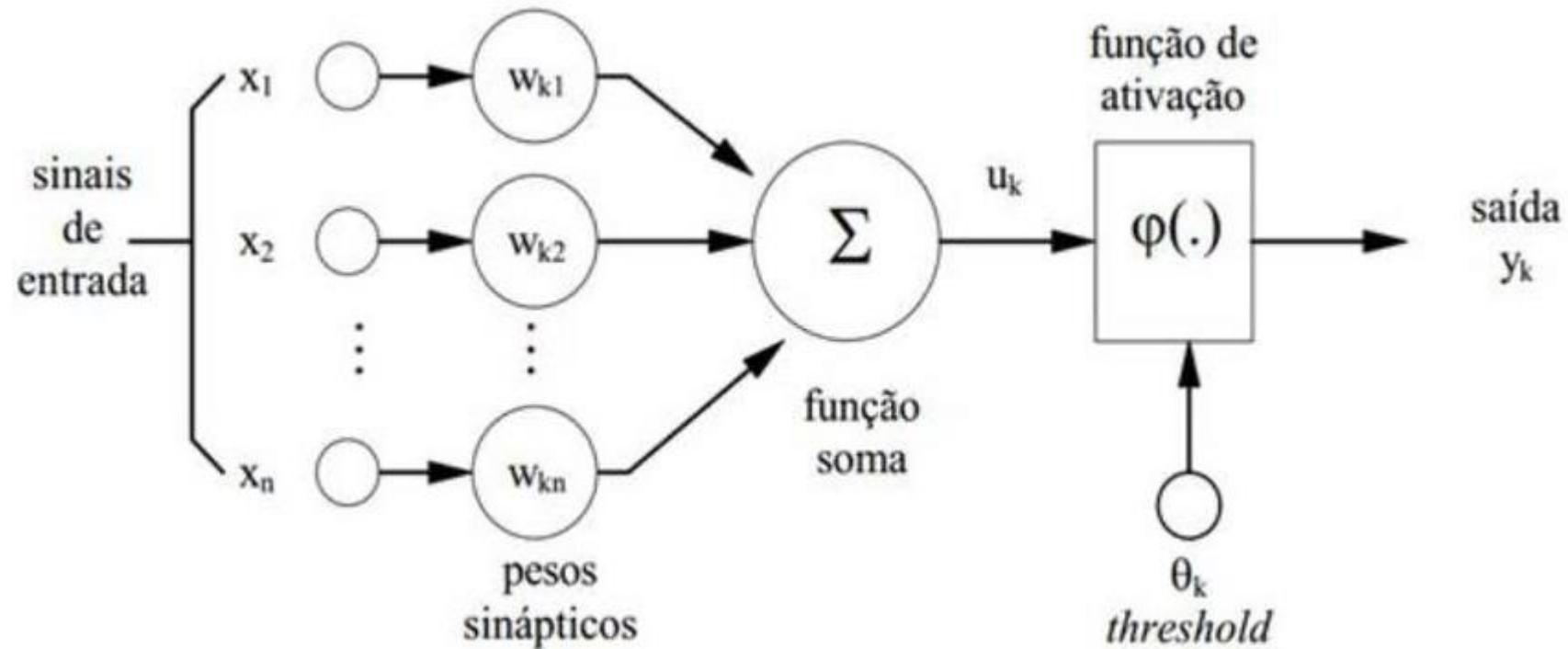
Ideação

- Tenta simular o funcionamento do cérebro humano
- É projetada para modelar a maneira como o cérebro realiza uma tarefa particular ou função de interesse

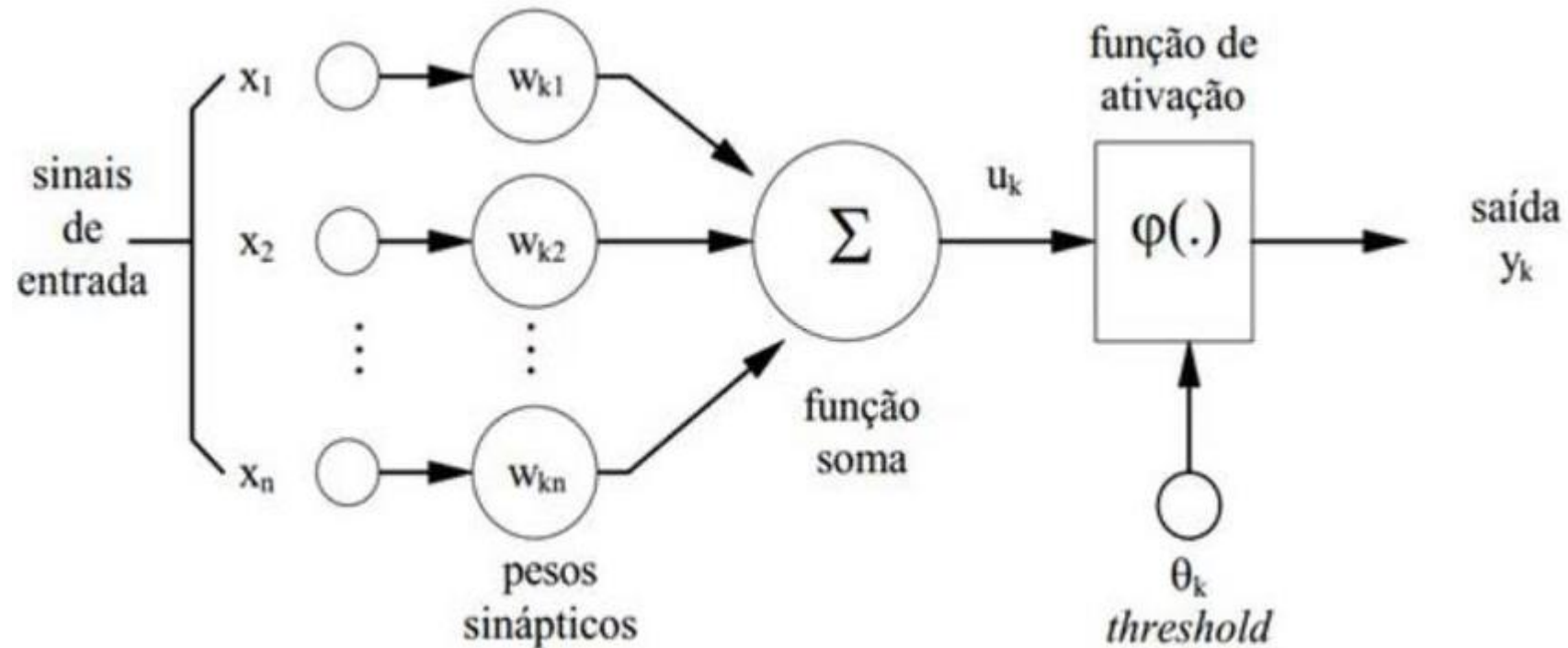
Ideação

- A RNA é um processador paralelamente distribuído constituído de unidades de processamento simples
- Tem a propensão natural para armazenar conhecimento e torná-lo disponível para o uso

Neurônio artificial

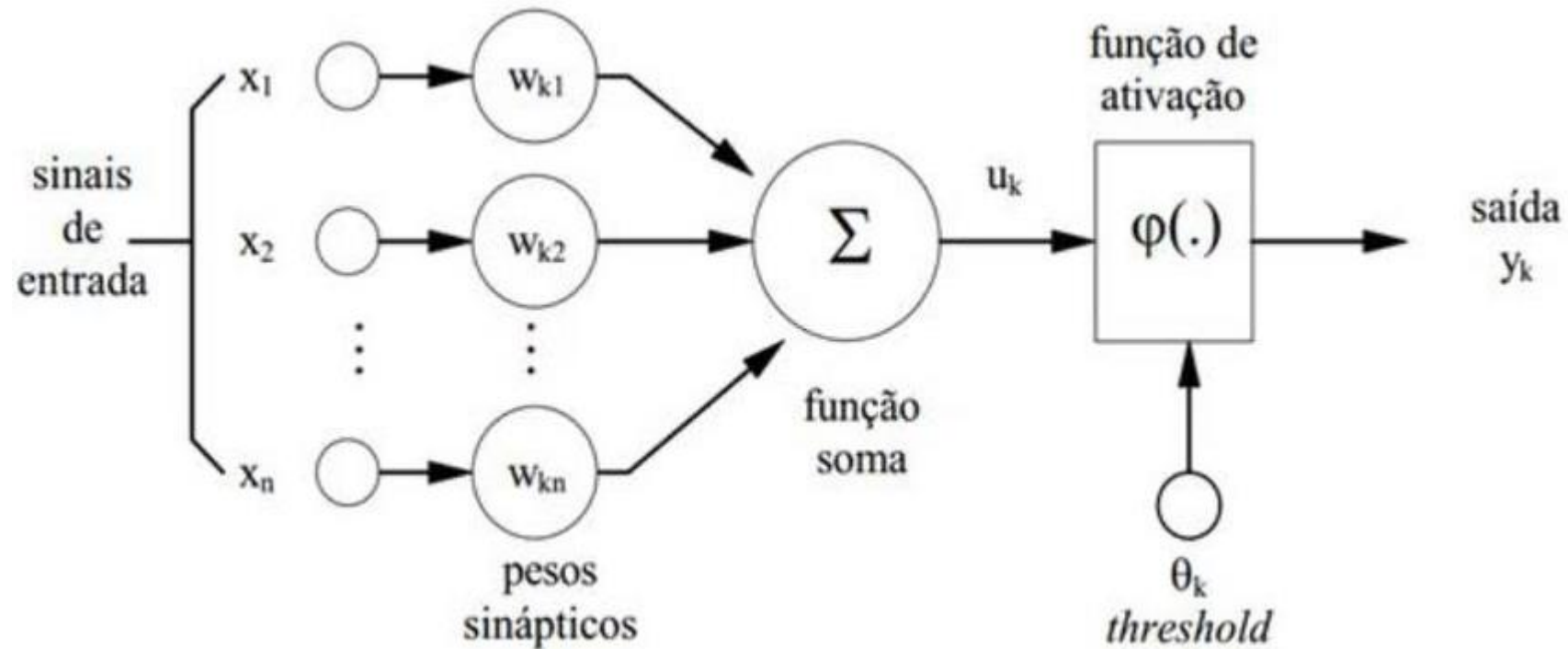


Neurônio artificial



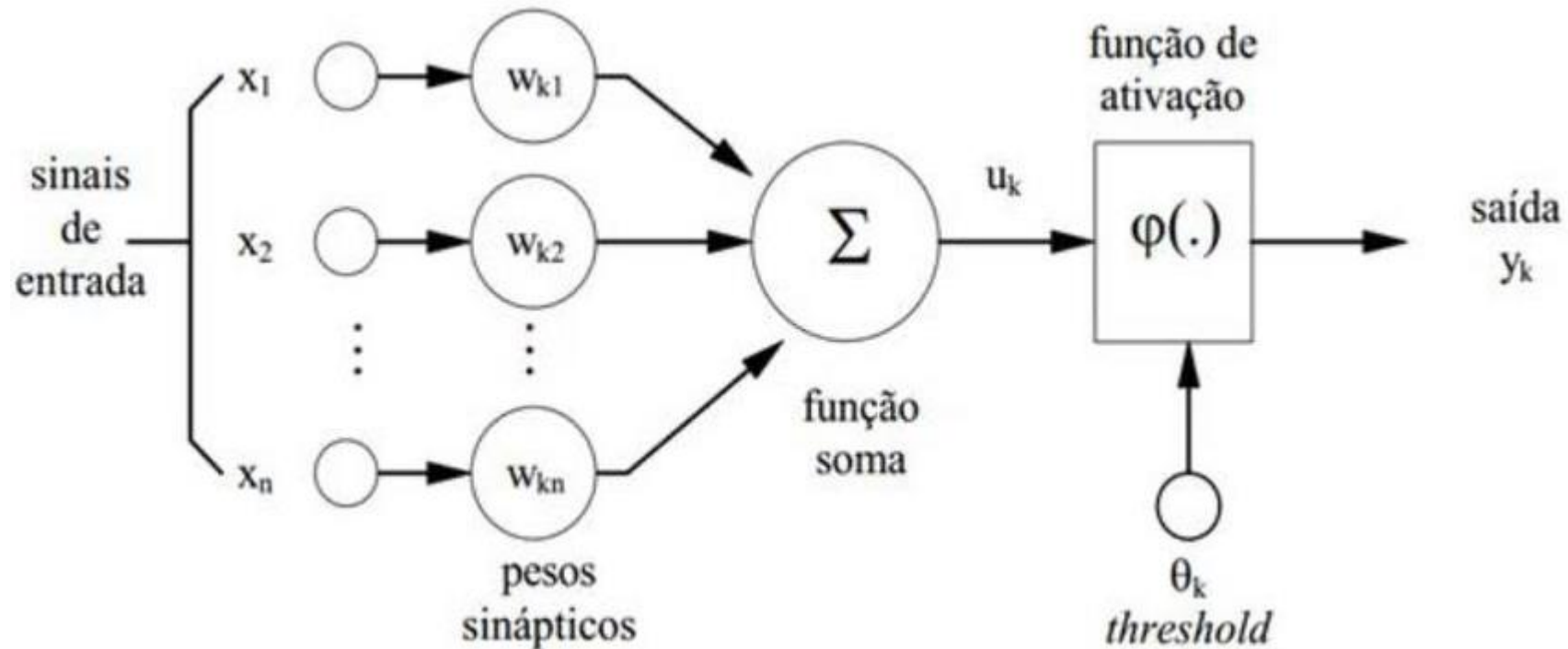
As entradas (x_1, x_2, \dots, x_n) são as características de um exemplo que queremos aprender ou estimar/classificar

Neurônio artificial



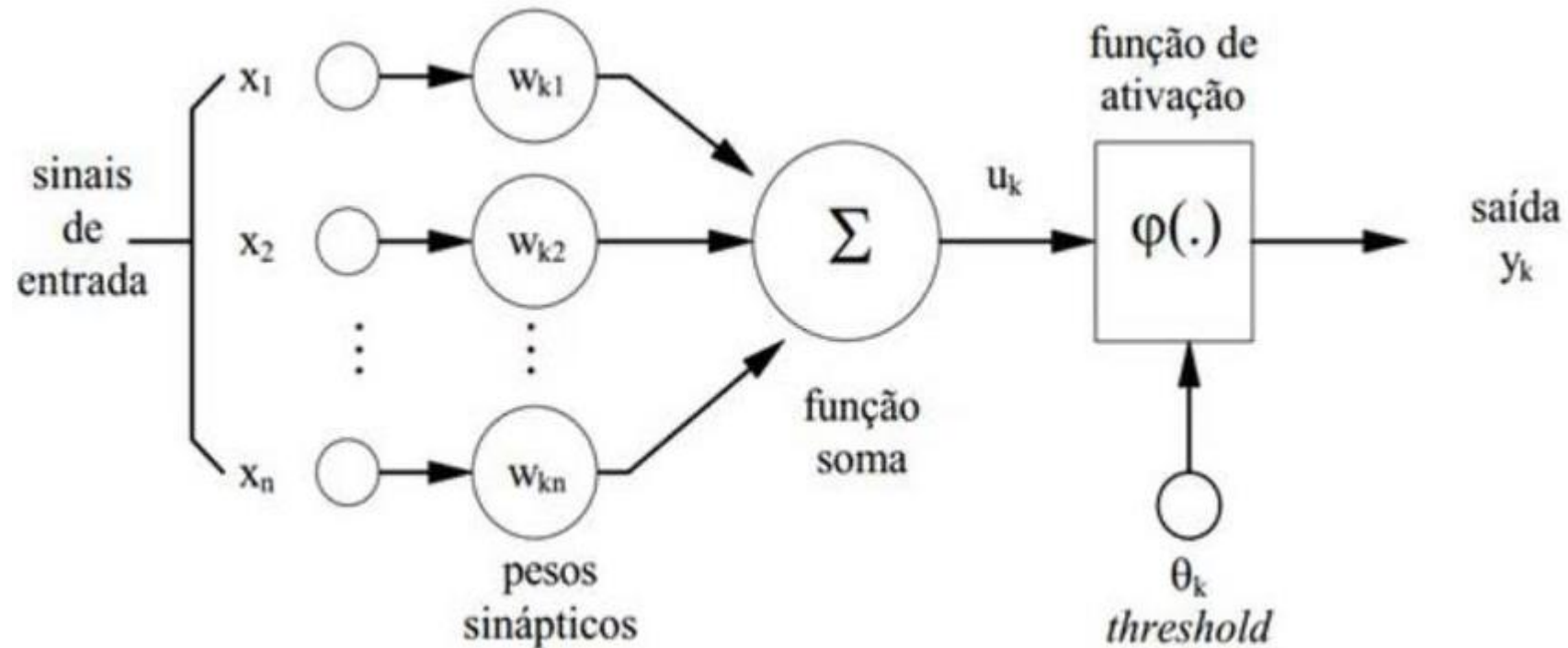
As entradas são multiplicadas por pesos sinápticos ($w_{k1}, w_{k2}, \dots, w_{kn}$)

Neurônio artificial



O resultado das multiplicações são combinadas em uma somatória, gerando um único resultado (u_k)

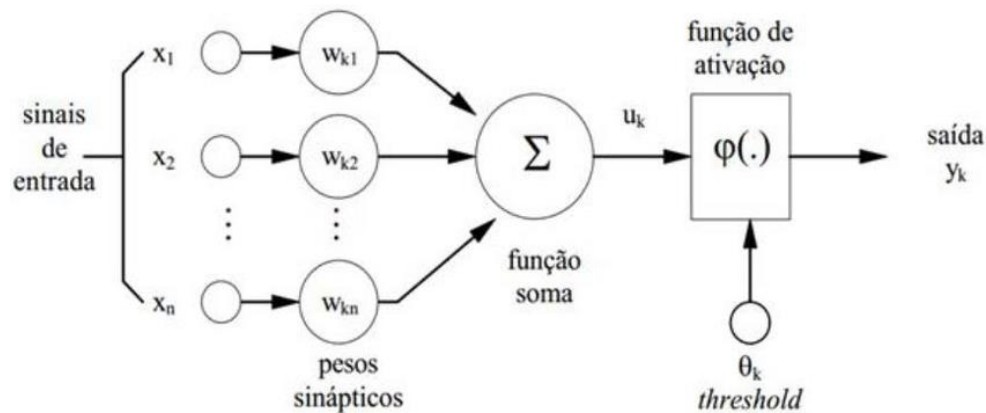
Neurônio artificial



Uma função de ativação em conjunto com um threshold (ou limiar) estabelece qual o sinal de saída desse neurônio (y_k)

Neurônio artificial

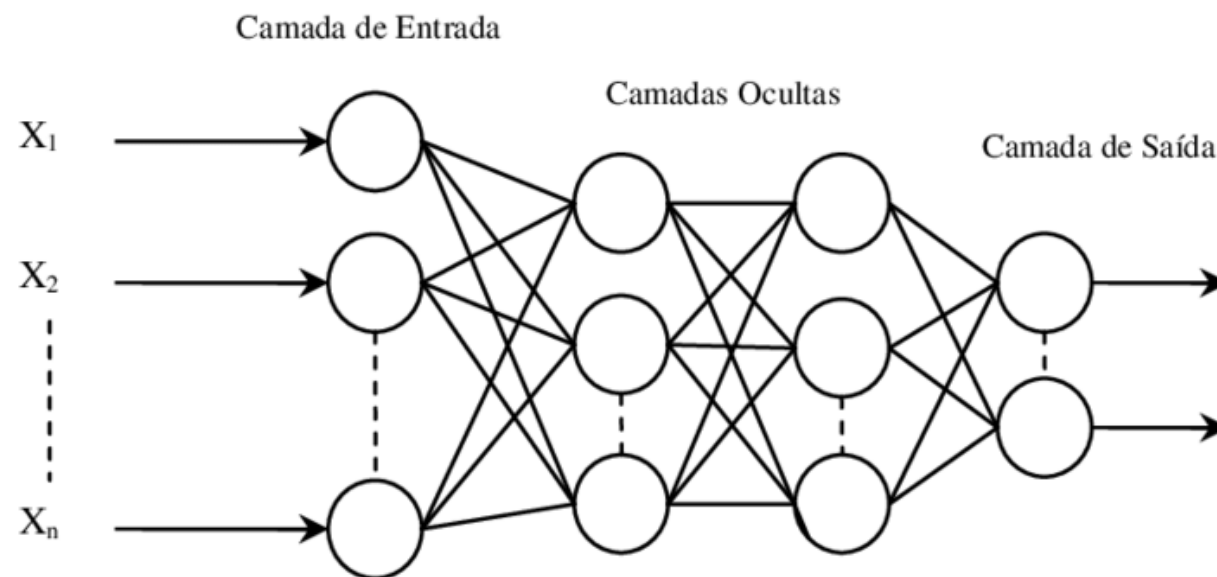
- Em termos matemáticos



$$u_k = \sum_{j=1}^m w_{kj} x_j$$
$$y_k = \varphi(u_k + b_k)$$

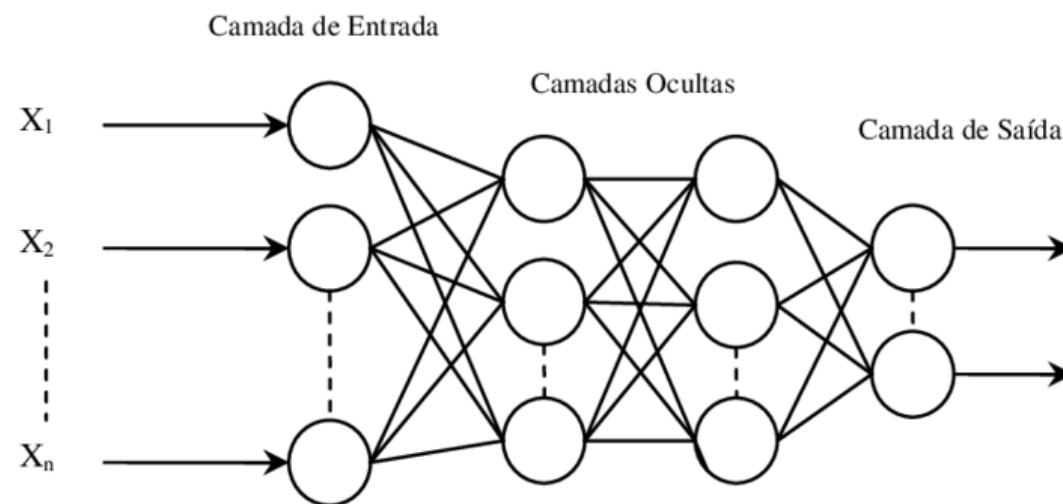
Rede neural artificial

- A saída de um neurônio pode ser a entrada do neurônio seguinte
- Combinando vários neurônios temos uma rede neural



Rede neural artificial

- O conjunto de neurônios podem formar camadas
- A camada de entrada recebe os dados de entrada
- As camadas ocultas (ou intermediárias) ajudam no processamento e detecção dos padrões
- A camada de saída é responsável por informar a resposta desejada da rede neural



Mas como a rede neural artificial aprende e consegue realizar estimativas corretas?

Processo de aprendizagem

- Aprendizagem é um processo pelo qual os parâmetros de uma rede neural são adaptados
 - Os pesos que ligam os dados de entrada à função de ativação devem ser atualizados de acordo com o resultado esperado, a fim de calibrar o modelo
- O tipo de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros (pesos sinápticos) ocorre
- Há diversos algoritmos de treinamento para realizar a aprendizagem

Processo de aprendizagem

- Passos:
 1. Definir os pesos iniciais
 2. Avaliar o desempenho da rede
 3. Corrigir os pesos quando necessário
 4. Repetir o passo 2 e 3 até que algum critério de parada seja atingido
- A definição dos pesos normalmente é feita de maneira aleatória, respeitando alguma distribuição
- A avaliação do desempenho da rede consiste em verificar se a rede está acertando os resultados esperados após processar as entradas com os pesos sinápticos atuais

Processo de aprendizagem

- Cada ciclo realizado é chamado de época
- O critério de parada pode ser
 - uma constante definindo o número de épocas a ser realizado
 - um grau de satisfação de acordo com valor obtido na avaliação de desempenho
 - um valor mínimo de mudança entre uma época e outra

Função de ativação

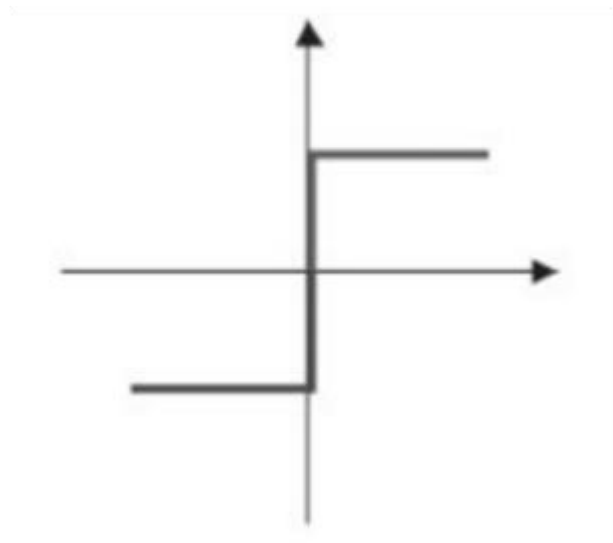
Função de ativação

- Serve para restringir a amplitude da saída de um neurônio
- O resultado da somatório pode ser um número muito grande ou muito pequeno, então é necessário padronizar a saída para que o processamento seguinte possa continuar sem problemas
- Existem várias funções de ativação

Tipos de função de ativação

- Limiar

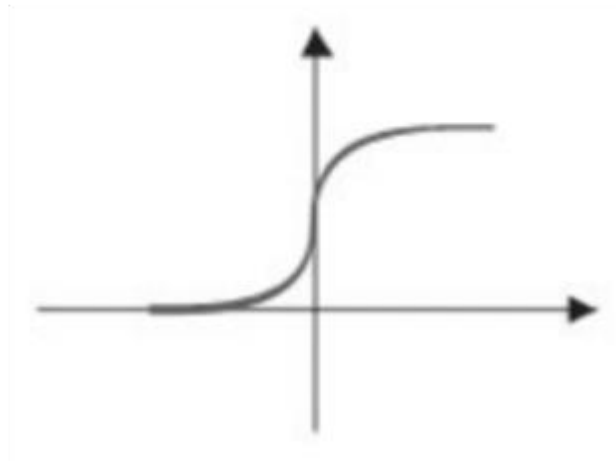
$$\varphi(v) = \begin{cases} 1 & \text{se } v \geq 0 \\ 0 & \text{se } v < 0 \end{cases}$$



Tipos de função de ativação

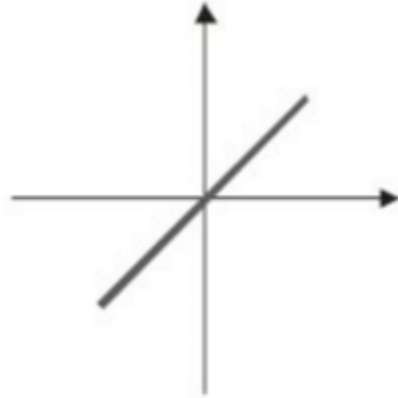
- Sigmoide

$$\varphi(v) = \frac{1}{1 + \exp(-av)}$$

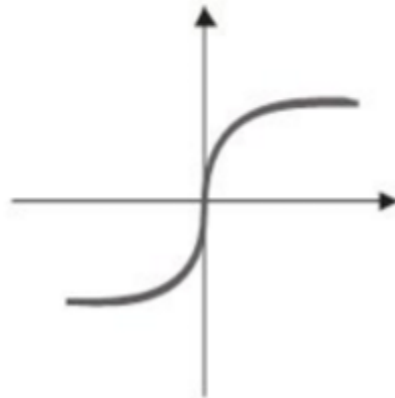


Tipos de função de ativação

- Linear

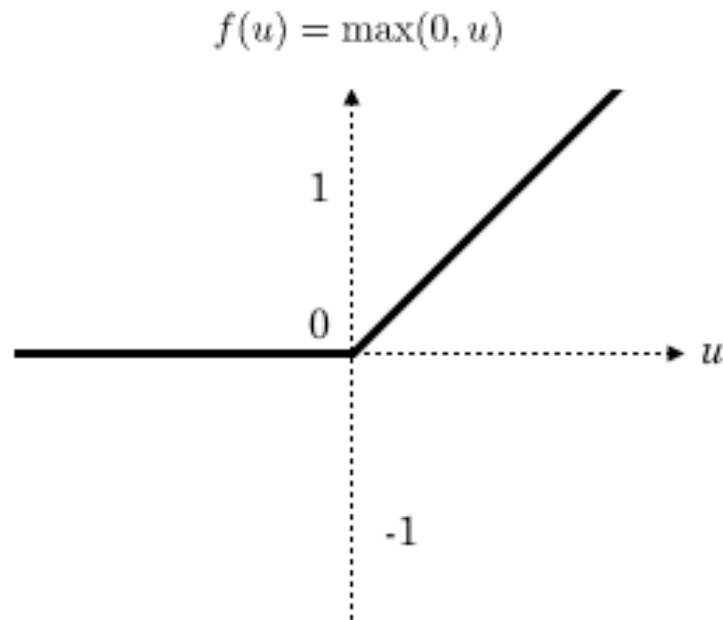


- Tangente hiperbólica



Tipos de função de ativação

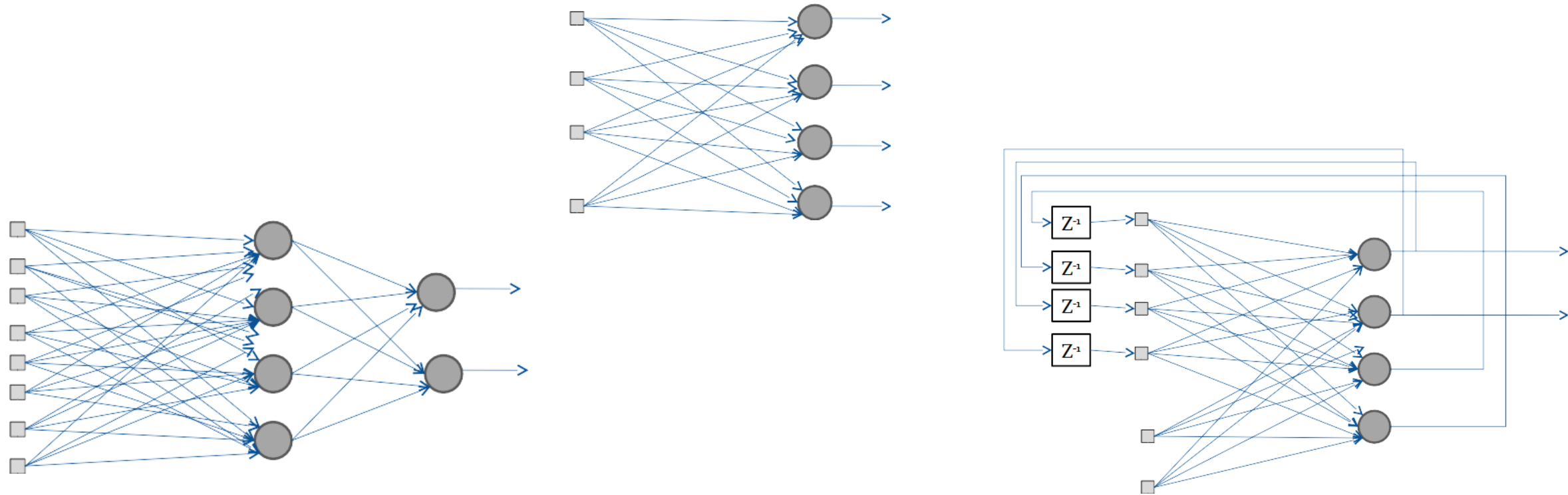
- Rectified linear unit (Relu)



Arquitetura da rede neural artificial

Arquitetura

- Uma rede neural artificial pode estar organizada de diversas maneiras



Número de camadas e neurônios

- Para a escolha do número de neurônios e camadas, deve ser considerado
 - Número de exemplos de treinamento
 - Quantidade de ruído presente nos exemplos
 - Complexidade da função a ser aprendida pela rede
 - Distribuição estatística dos dados de treinamento

Número de camadas e neurônios

- Quanto mais neurônios e/ou camadas:
 - Maior a possibilidade de resolver problemas complexos
 - Maior tempo de treinamento
 - Maior possibilidade de overfitting

Exercício 1

- Explore esta aplicação
 - <https://phiresky.github.io/neural-network-demo/>
- Modifique os pontos (dados)
- Avalie o desempenho com poucas e muitas camadas, e com poucos e muitos neurônios