



Introdução Inteligência Computacional

Profa. Dra. Ana Paula Abrantes de Castro e Shiguemori

anapaula.acs@ifsp.edu.br

Instituto Federal de Educação, Ciência e Tecnologia – IFSP

Dr. Elcio Hideiti Shiguemori

elcio@ieav.cta.br

Instituto de Estudos Avançados - IEAv

Curso: Análise e Desenvolvimento de Sistemas
4º. Semestre - IIC14 – 4 aulas Semanais



- As redes neurais artificiais são modelos baseados no funcionamento das redes dos neurônios biológicos.
- Estes modelos têm sido utilizados na solução de uma grande variedade de problemas em diversas áreas.
- Algumas de suas características, como baixo custo computacional, tolerância a falhas e capacidade de generalização, as tornam atrativas para implementações em software e hardware.

Redes Neurais Artificiais

Disciplina tecnológica que trata dos sistemas de processamento de informação paralelos, distribuídos e adaptáveis, que autonomamente desenvolvem capacidades de processamento da informação como respostas ao ambiente de informação.

(Hecht-Nielsen, 1990).

- Não necessita do desenvolvimento de regras ou de algoritmos e reduz a quantidade de software que deve ser desenvolvido.
- Estruturas Primárias de Processamento da Informação de Interesse: Redes Neurais

Histórico

1943

McCulloch e Pitts - "A Logical Calculus of the Ideas Immanent in Nervous Activity." Bulletin of Mathematical Biophysics, 1943, 5:115-133.

1949

Donald **Hebb** - Postulado para aprendizagem (Regra de Hebb)

1958

Frank **Rosenblatt** - **Perceptron** (classificador de padrões separáveis linearmente)

1969

Minsky e Paper - "Perceptron: an introduction to computational geometry". MIT Press, Massachusetts, 1969.
Perceptron não resolvem problemas não lineares.

Histórico

Anos 70

Poucas pesquisas (***Fukushima, Grossber, Kohonen***)

1982

John ***Hopfield*** - Paper enfatizando as propriedades associativas das RNA - Reiniciaram as pesquisas - relação entre redes recorrentes e sistemas físicos

1986

D.E. ***Rumelhart***, G.E. ***Hinton*** e R.J. Williams, “Learning representations by ***back-propagating*** errors”, Nature, 323:533-536, 1986.

Retomadas pelas Pesquisas

Redes Neurais Artificiais - RNA

- São técnicas computacionais que apresentam um modelo matemático inspirado no funcionamento do cérebro humano
- Adquirem conhecimento através da experiência
- Modelam a forma do cérebro desempenhar tarefas ou funções
- Implementação em *hardware* ou *simuladas em software*

Redes Neurais Artificiais - RNA

- Interconexão maciça das células computacionais (*neurônios, ou unidades de processamento*) para alcançar um bom desempenho
- Desempenham as computações através de um processo de aprendizagem
- Processador paralelo distribuído com capacidade natural para armazenar conhecimento experimental para uso posterior.

Definição de uma RNA

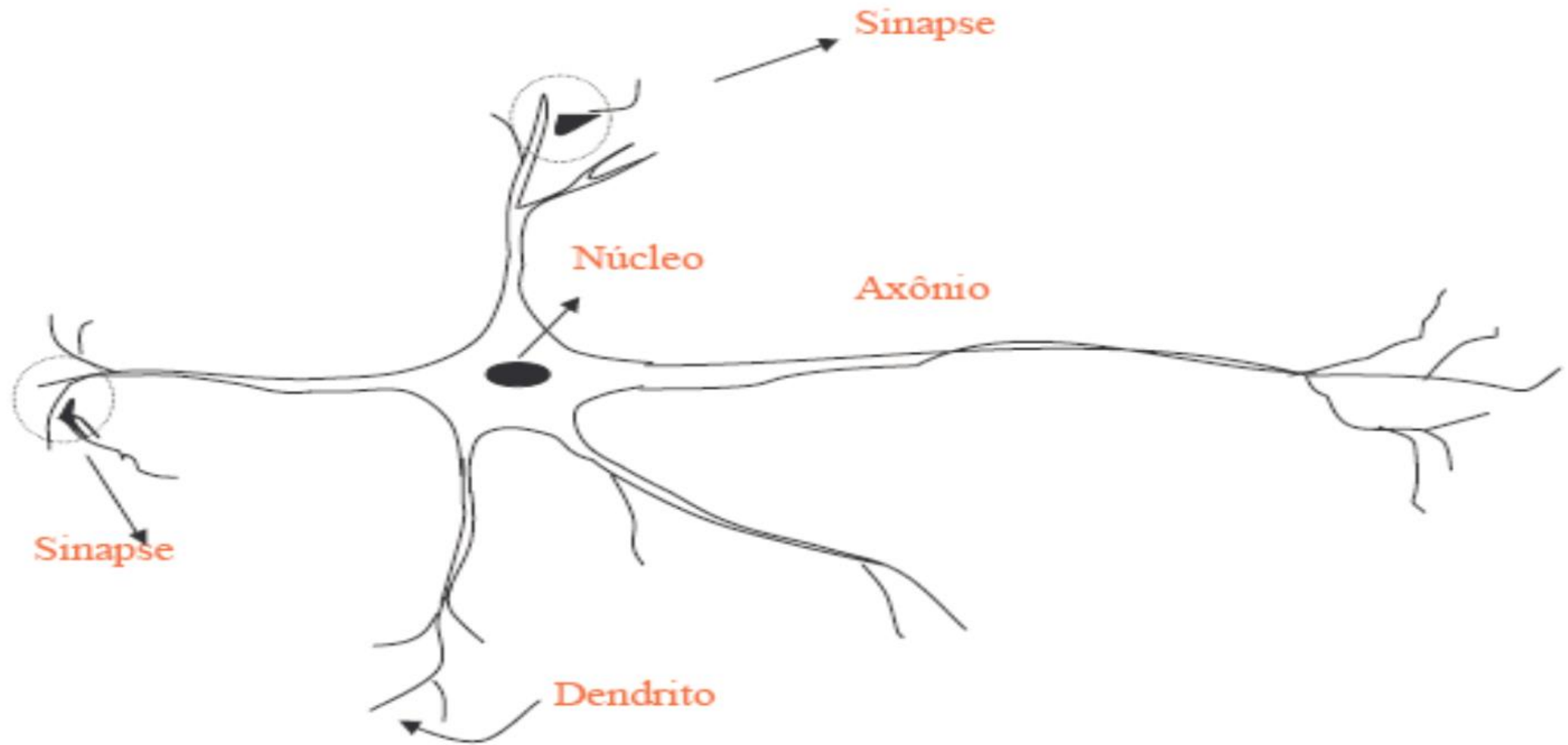
- ▶ Imita o cérebro de duas formas:
 - ▶ Adquire conhecimento através de aprendizagem.
 - ▶ Armazena conhecimento nas conexões entre neurônios (**pesos**).
- ▶ A aprendizagem é feita por um *algoritmo* que modifica os pesos da rede de uma forma ordenada para adquirir uma arquitetura previamente desejada
- ▶ Além da modificação dos pesos sinápticos uma RNA também pode modificar a própria *topologia*

Motivação das Redes Neurais



- Difere dos computadores digitais (na forma de fazer computação).
- Estrutura básica constituinte - neurônio (Ramón e Cajál, 1911 -Haykin, 1994).
- Neurônio execução *milisegundos* ($10^{-3}s$) X porta lógica *nanosegundos* ($10^{-9}s$)
- No. de neurônios e interconexões supera a baixa velocidade de operação (10 bilhões de neurônios e 60 trilhões de conexões -Haykin, 1994).
- Sistema complexo, não linear e paralelo.

Componentes de um neurônio



Componentes de um Neurônio

Axônio: transmissão de sinais a partir do corpo celular poucas ramificações e compridos

Dendritos: conduzem sinais para a célula; têm muitas ramificações.
(zonas receptivas)

Sinapses: Unidades estruturais e funcionais elementares que medeiam as conexões entre os neurônios

Definição de uma RNA

Aprendizagem

- *Algoritmo* modifica os pesos da rede de uma forma ordenada para adquirir uma arquitetura previamente desejada
- Também pode modificar a própria *topologia*

Benefícios

- Não linearidade
- Mapeamento de entrada/saída
- Adaptabilidade
- Informação contextual
- Tolerância à falha
- Implementação em VLSI
- Analogia neurobiológica

Funcionamento

Aprendizagem:

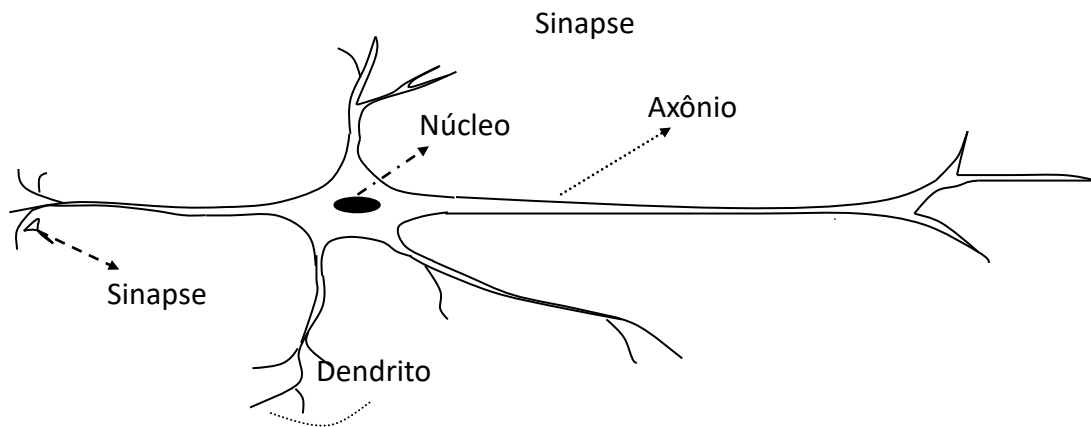
Consiste no processo de Adaptação dos pesos sinápticos das conexões e os níveis de bias dos neurônios em resposta as entradas.

Ativação:

Consiste no processo de receber uma entrada e produzir uma saída com os pesos e bias obtidos na fase de aprendizagem.

Inspiração Biológica

Neurônio Biológico



O neurônio Biológico pode ser visto como sendo o dispositivo computacional elementar básico do sistema nervoso.

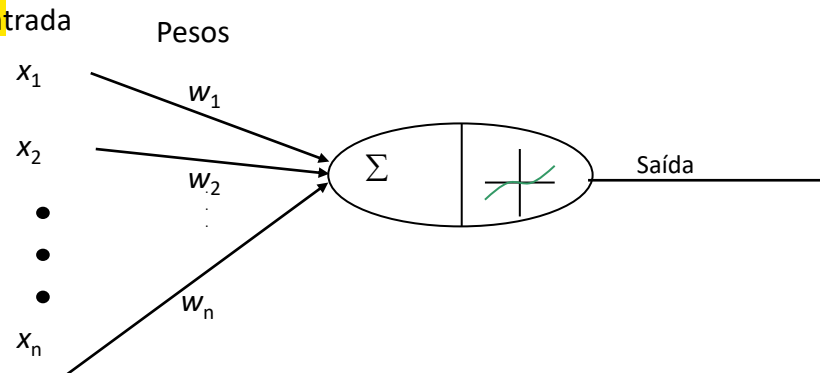
- Possui muitas entradas e uma saída.
- Entradas → conexões sinápticas
- Saída → axônio

O estímulo que chega à sinapse é transferida à membrana dendrital que dá origem a uma conexão excitatória ou inibitória.

Assim como os neurônios biológicos, os neurônios artificiais têm inúmeras entradas dadas pelos níveis de estímulos.

- Cada entrada é multiplicada por um peso sináptico.
- O resultado da multiplicação é somado
- E então passada por uma função de ativação

Neurônio Artificial

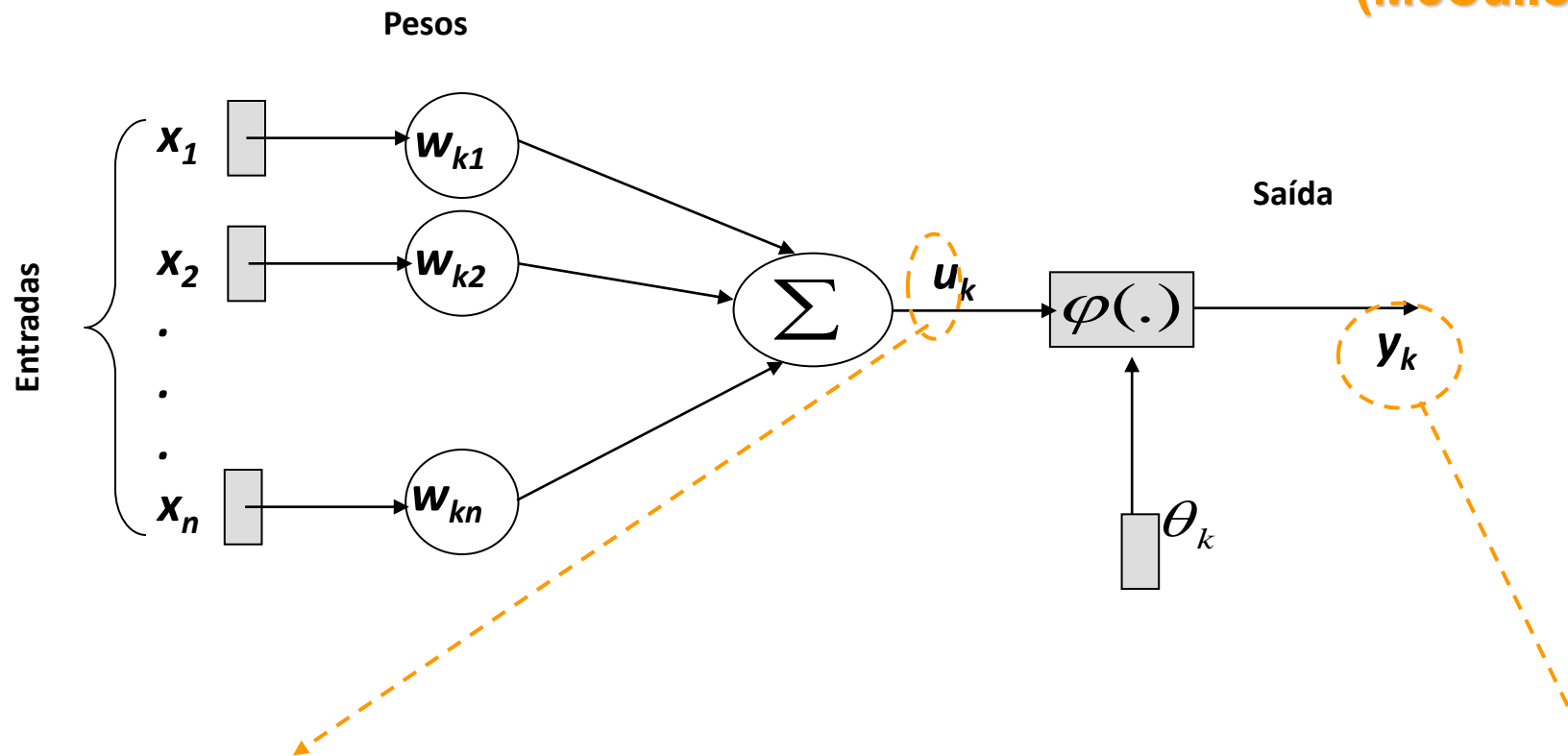


$$v_j = \sum_{k=1}^n w_{jk} x_k + b_j$$

$$y_j = \varphi(v_j)$$

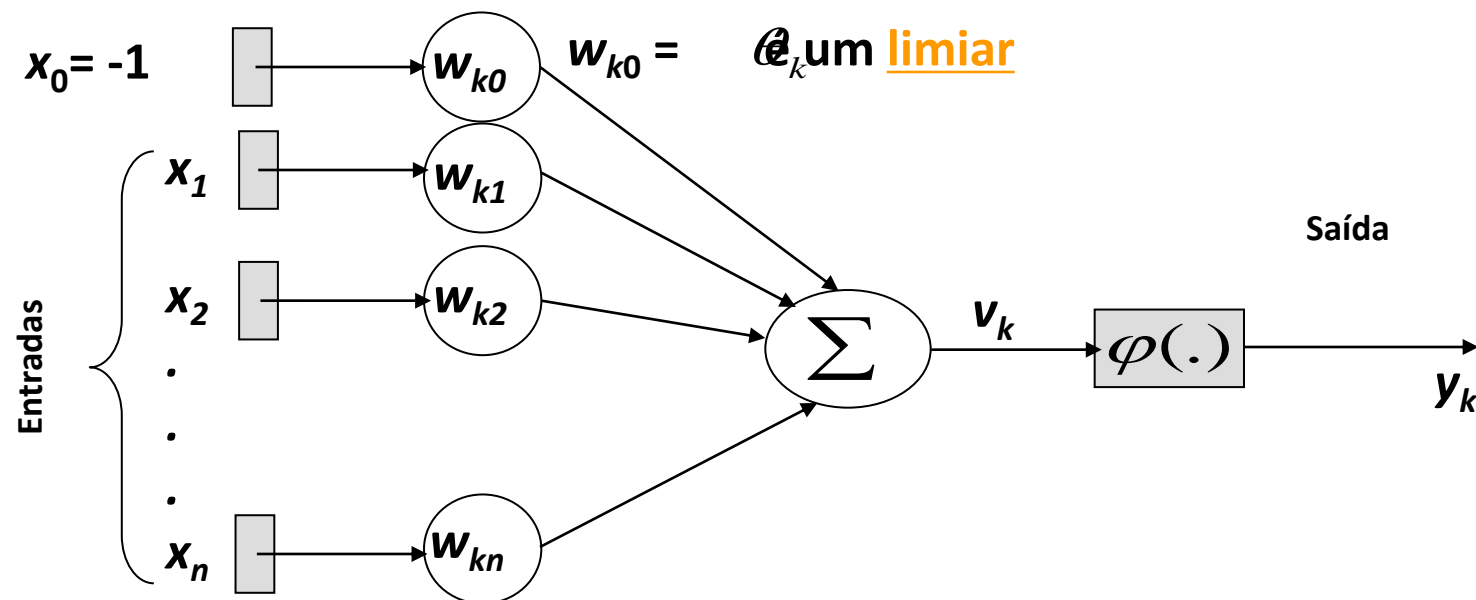
Modelo do Neurônio Artificial

(McCulloch e Pitts, 1943)

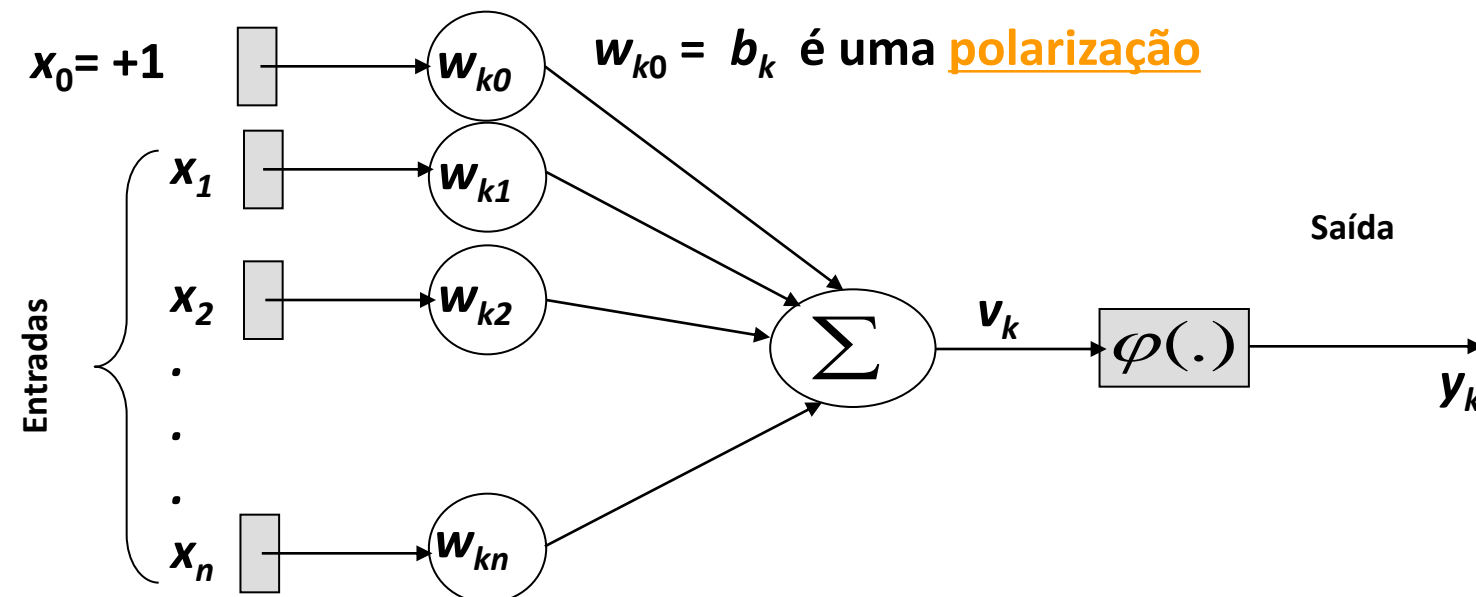


$$\textcircled{u}_k = \sum_{j=1}^n w_{kj} x_j \quad \longrightarrow \quad v_k = u_k - \theta_k \quad \longrightarrow \quad \textcircled{y}_k = \varphi(v_k)$$

Modelo do Neurônio Artificial



$$v_k = \sum_{j=0}^n w_{kj} x_j$$

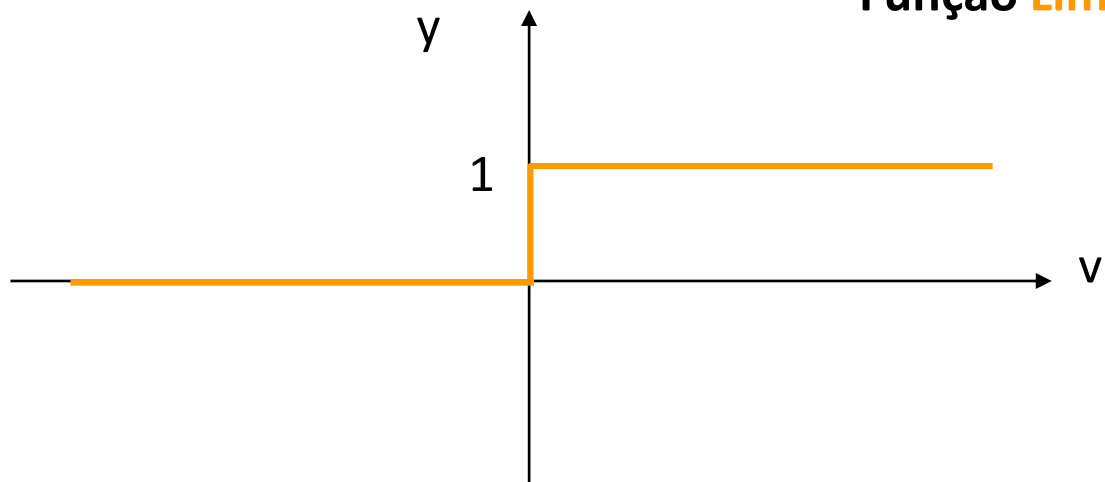


$$y_k = \varphi(v_k)$$

Funções de Ativação

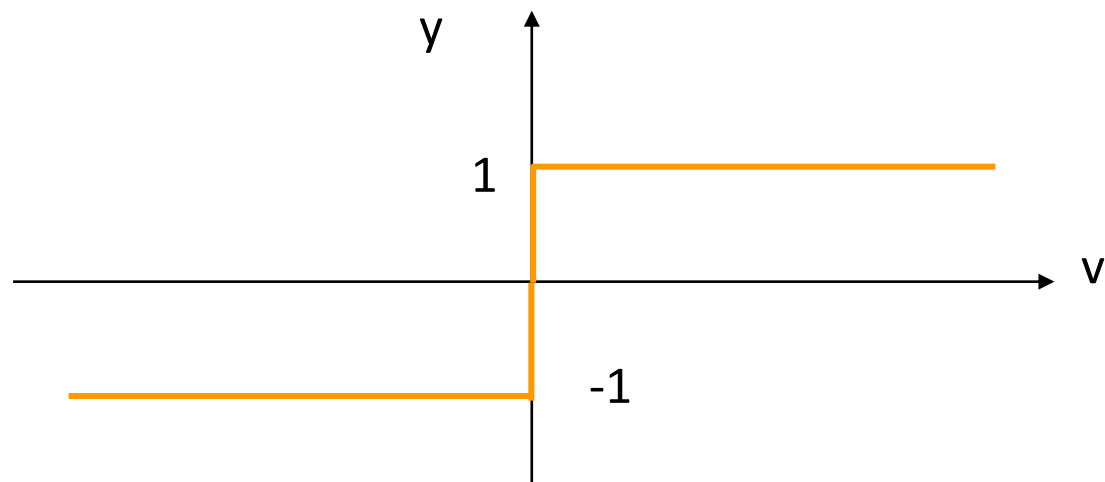
$$y = \varphi(v)$$

Função Limiar



$$y = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases}$$

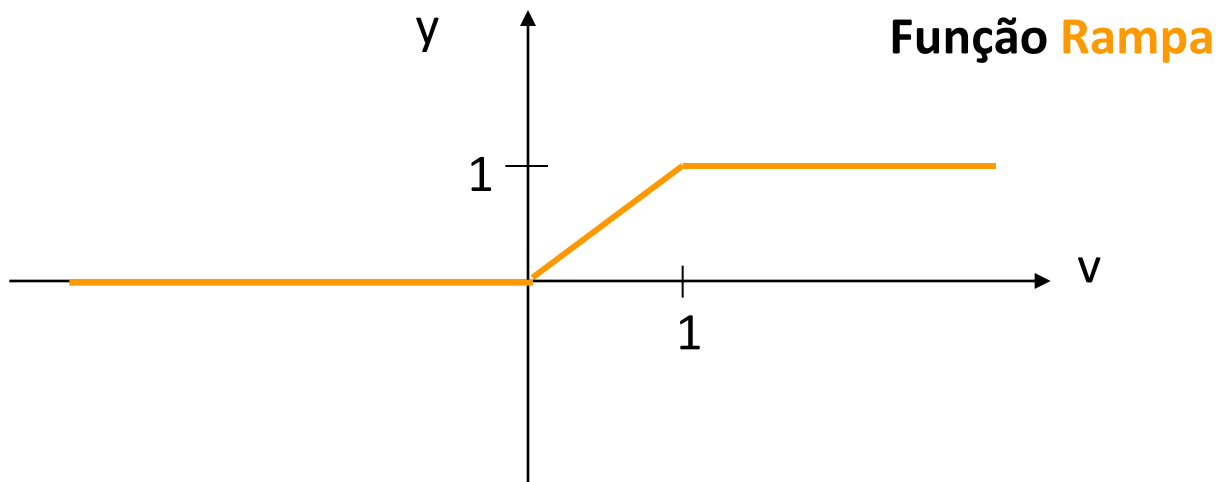
Função Sinal



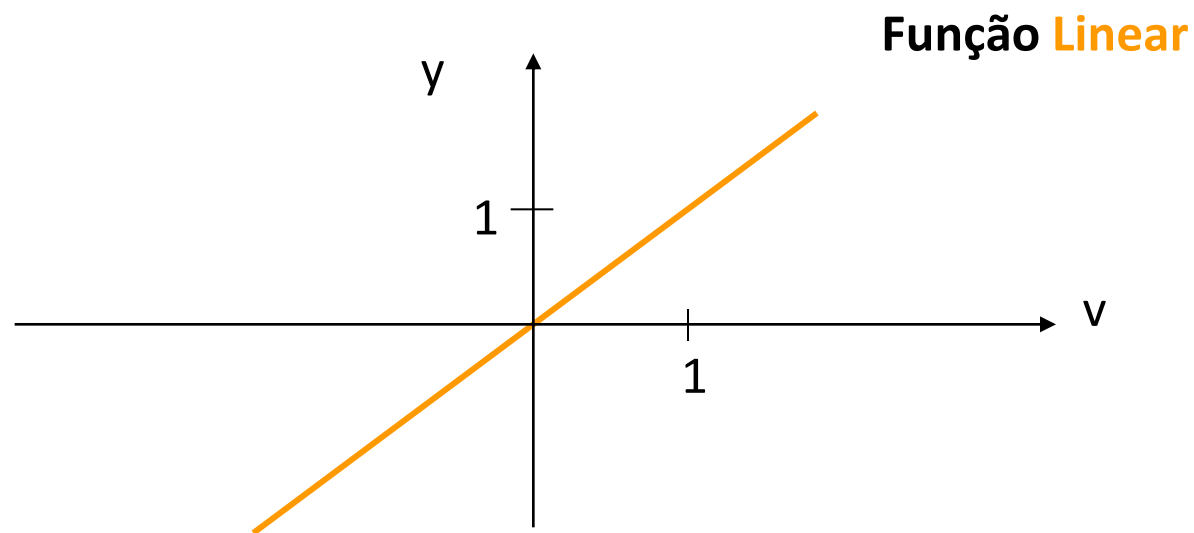
$$y = \begin{cases} +1, & v \geq 0 \\ -1, & v < 0 \end{cases}$$

Funções de Ativação

$$y = \varphi(v)$$



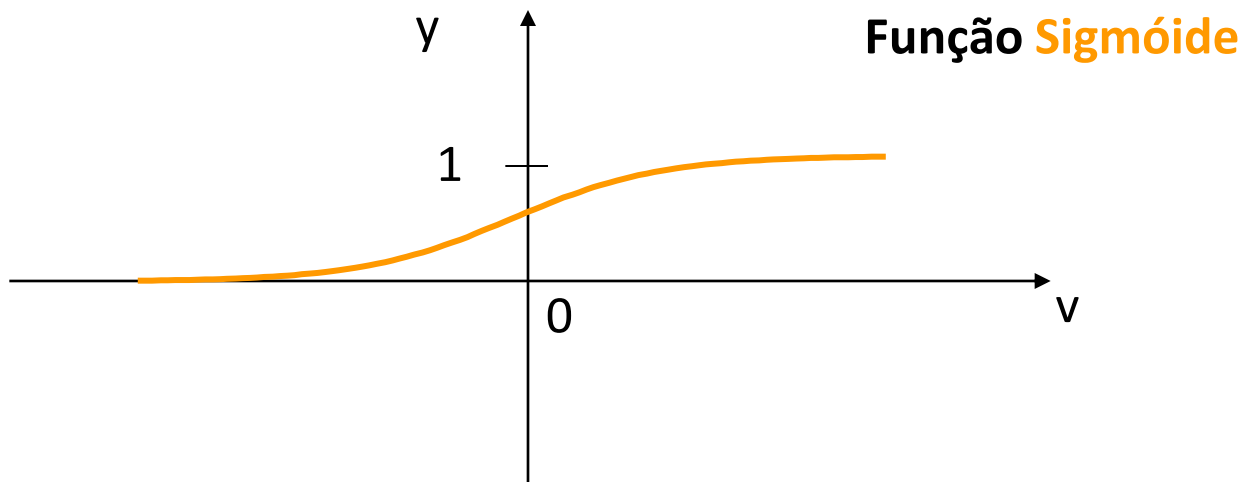
$$y = \begin{cases} 0, & v < 0 \\ v, & 0 \leq v < 1 \\ 1, & v > 1 \end{cases}$$



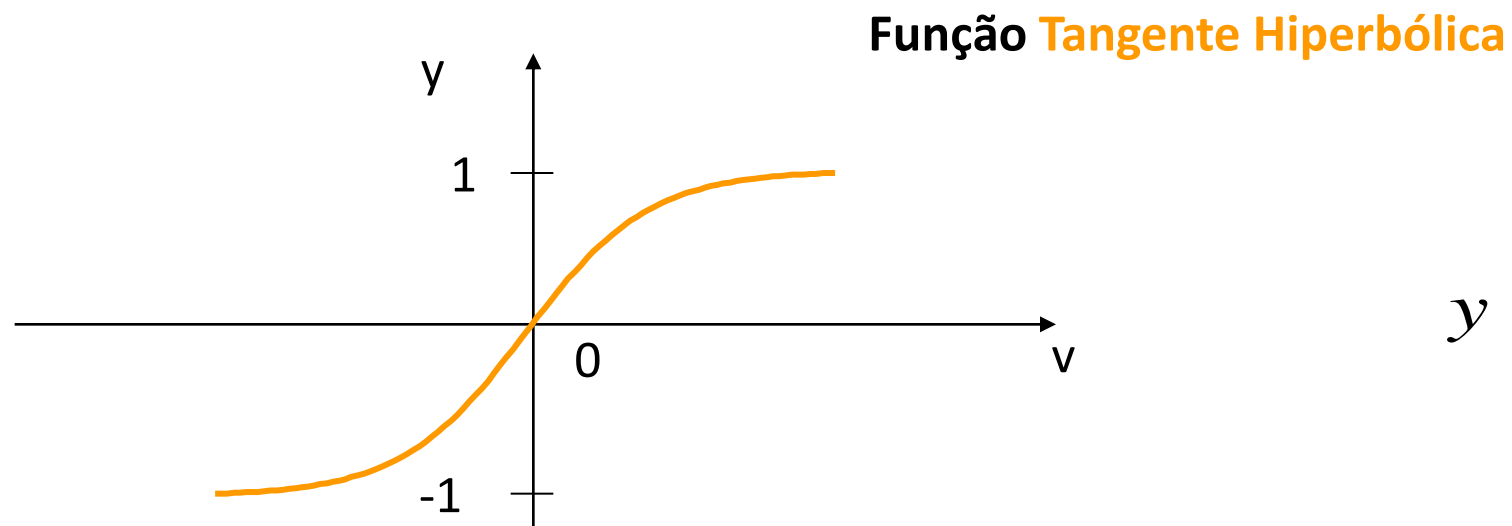
$$y = v$$

Funções de Ativação

$$y = \varphi(v)$$



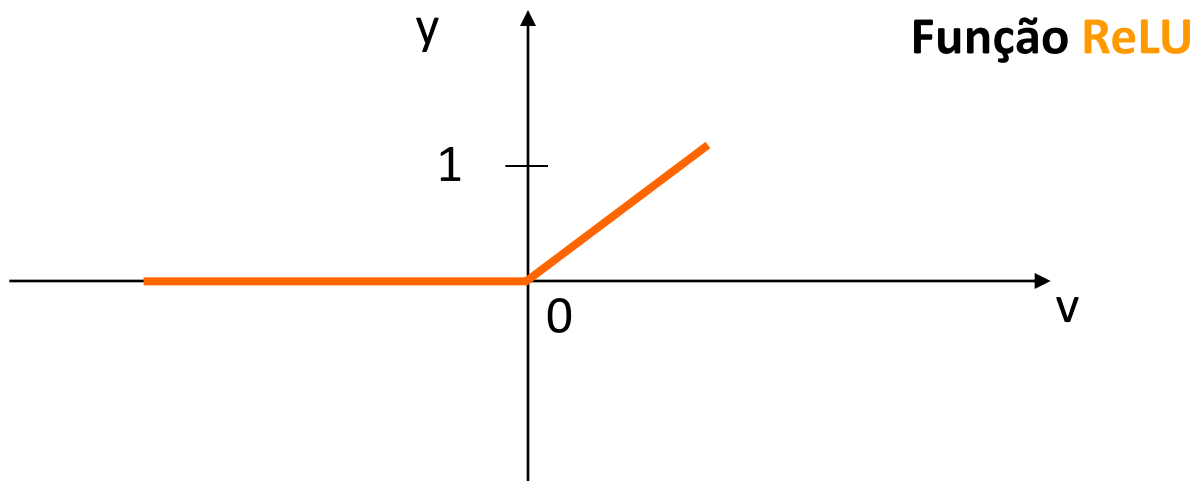
$$y = \frac{1}{(1 + e^{-v})}$$



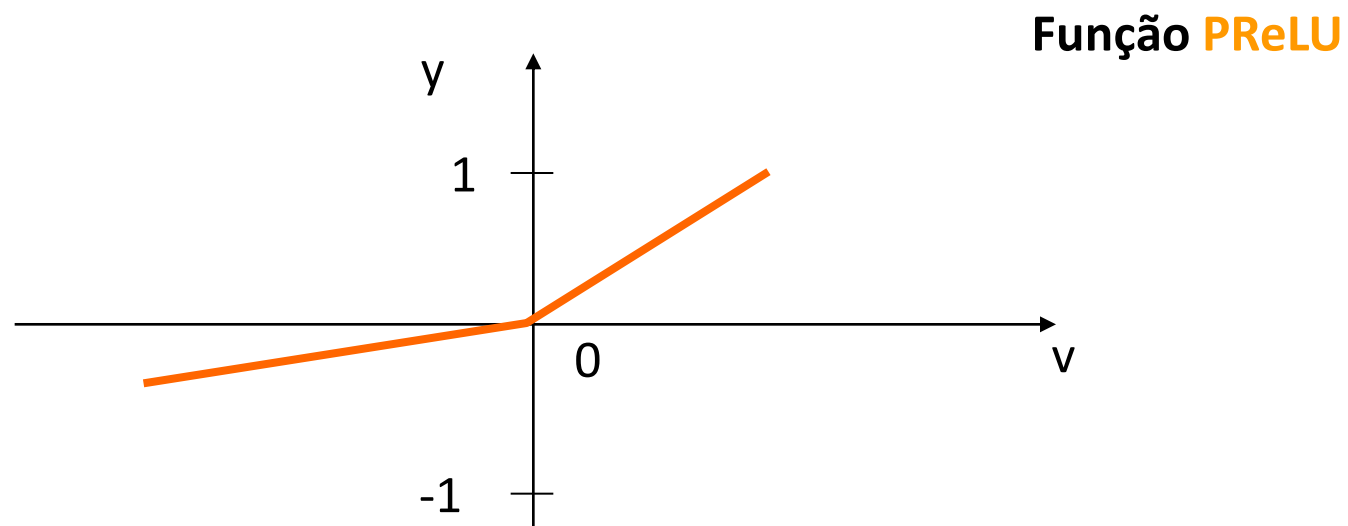
$$y = \tanh\left(\frac{v}{2}\right) = \frac{1 - e^{-v}}{(1 + e^{-v})}$$

Funções de Ativação

$$y = \varphi(v)$$



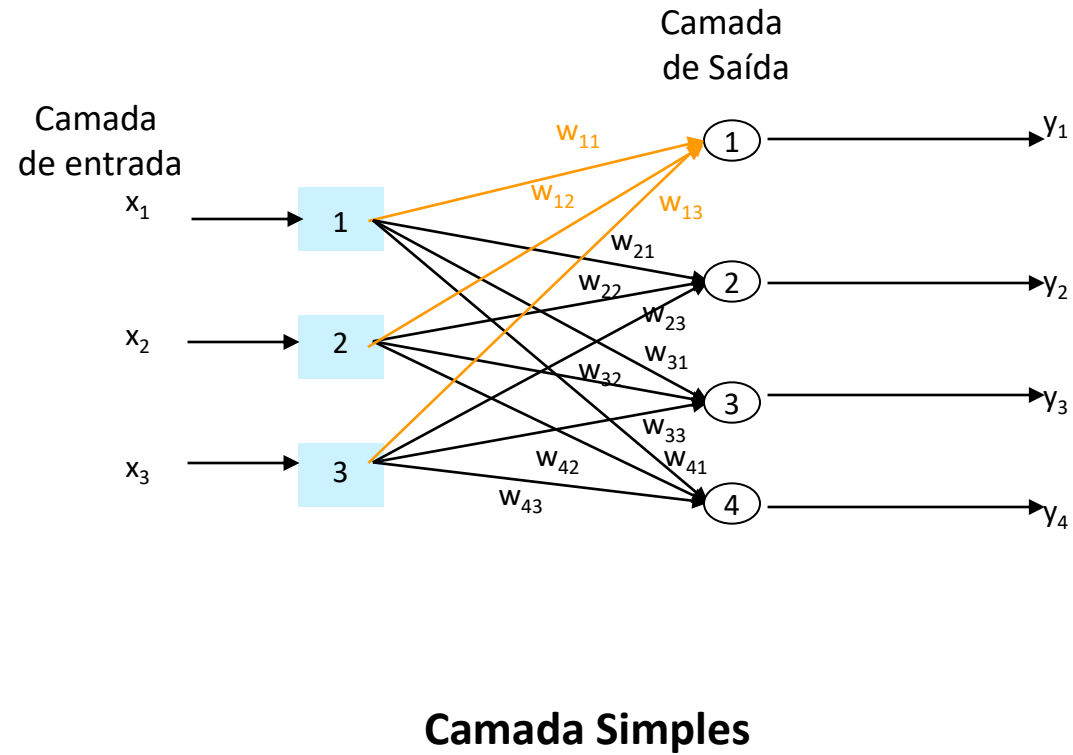
$$y = \max[0, v]$$



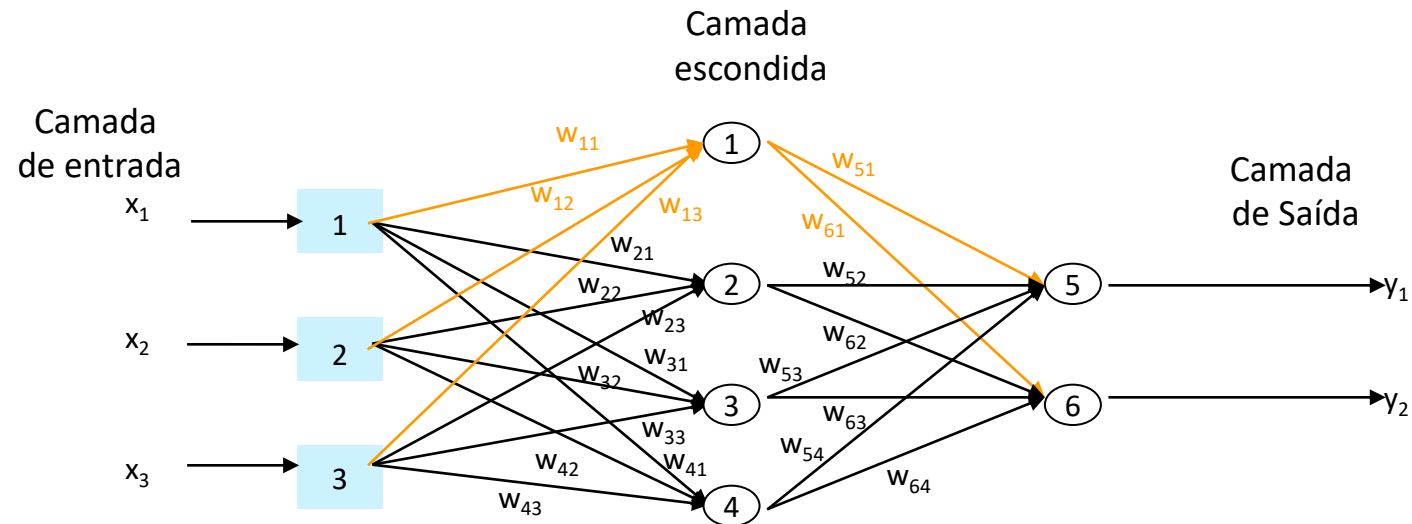
$$y = \max[av, v], a = 0,1$$

Modelos de Arquiteturas

Rede FeedForward

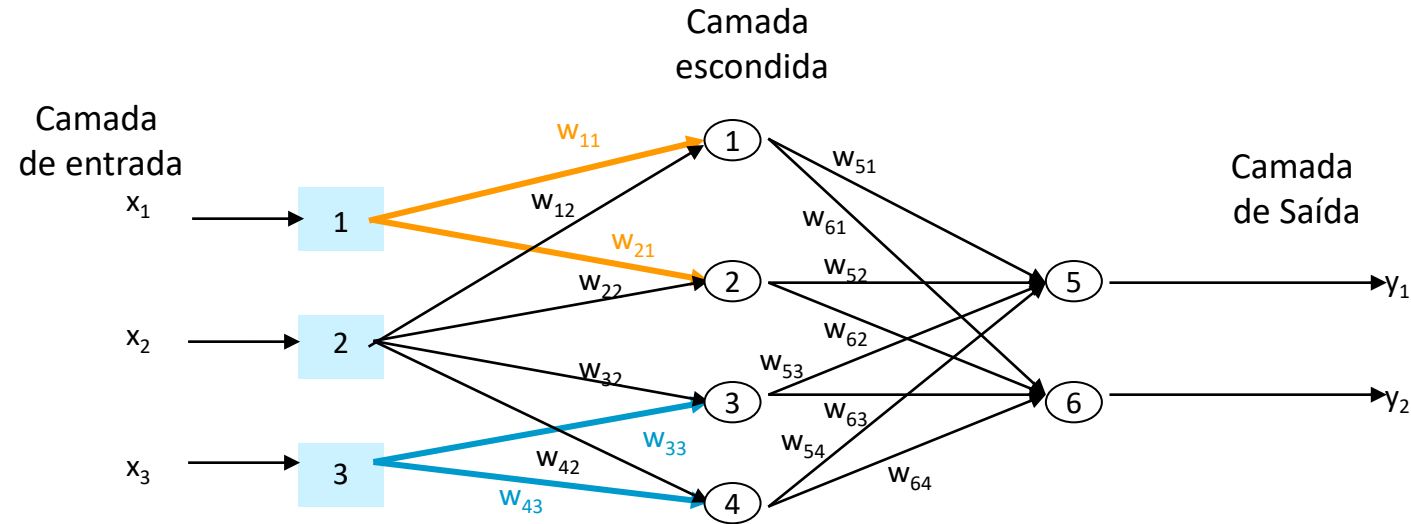


Rede FeedForward



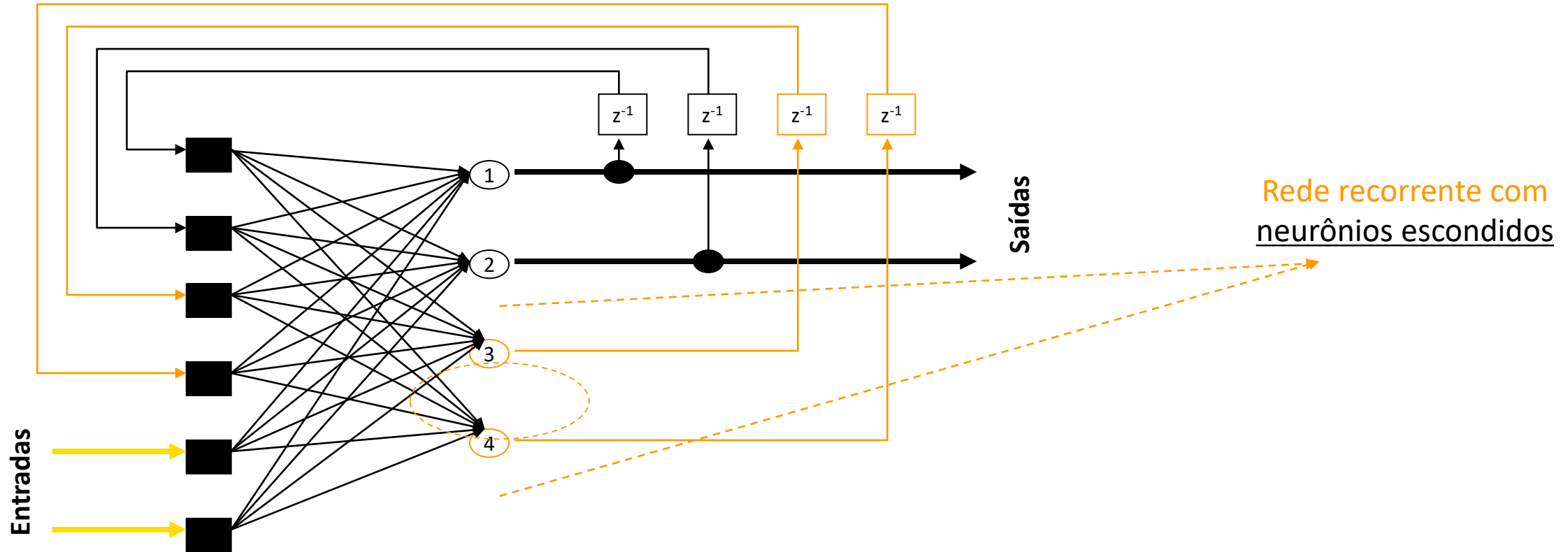
Múltiplas Camadas Totalmente conectadas

Rede FeedFoward

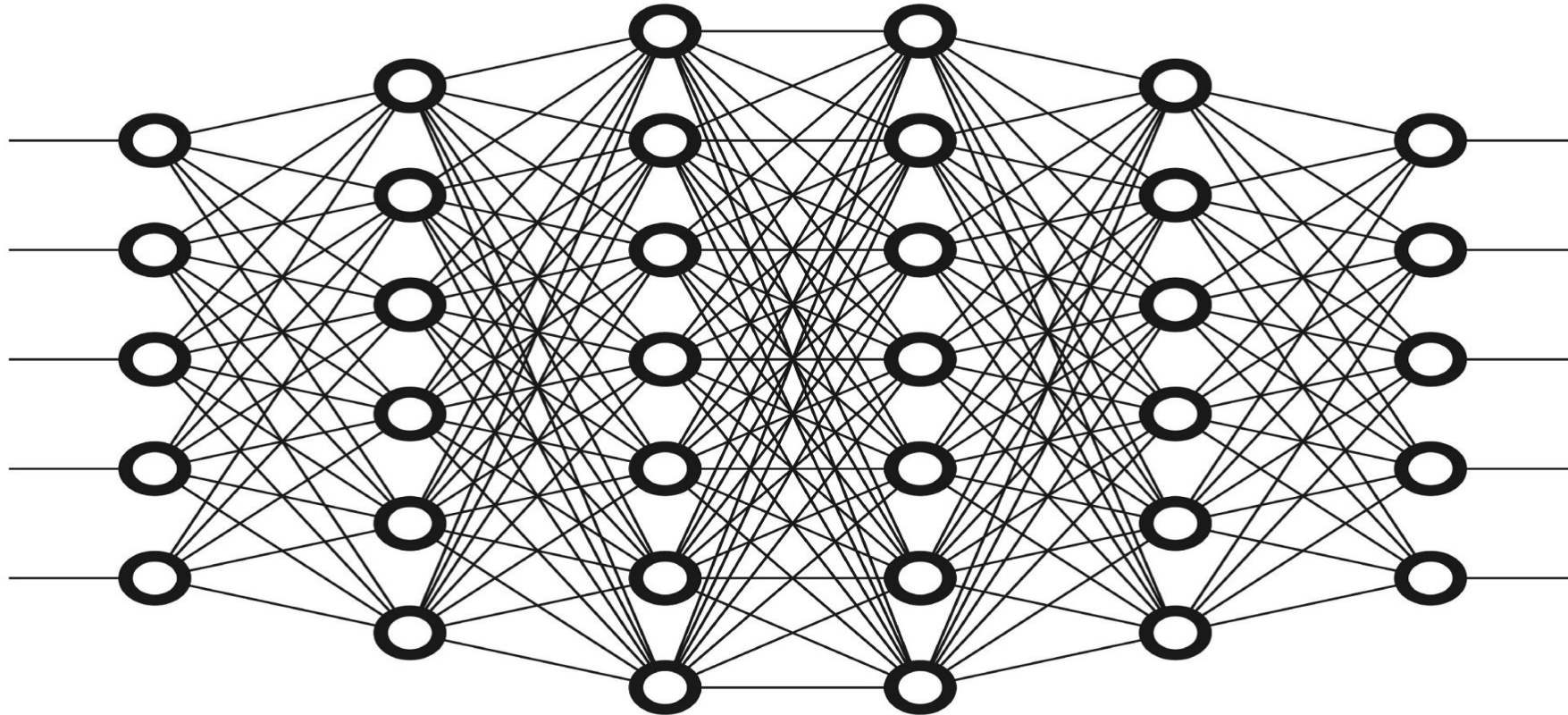


Parcialmente conectada

Rede FeedFoward



Rede Profunda





Dúvidas?



Referências Bibliográficas

- José Demísio Simões da Silva – Notas de Aula
- Ana Paula A. C. Shiguemori – Notas de Aula
- Elcio Hideiti Shiguemori – Notas de Aula