UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

Inteligência Computacional

CP78D

Redes Neurais

Aula 8

Prof. Daniel Cavalcanti Jeronymo

Universidade Tecnológica Federal do Paraná (UTFPR)
Engenharia Eletrônica – 7º Período

Plano de Aula

- Preditor Linear
- Perceptron
- Multilayer Perceptron (MLP)

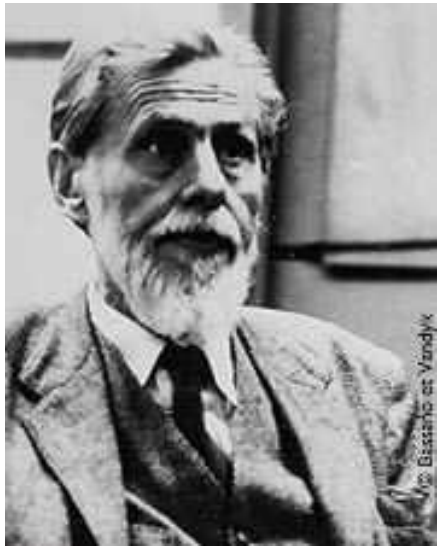
Perceptron

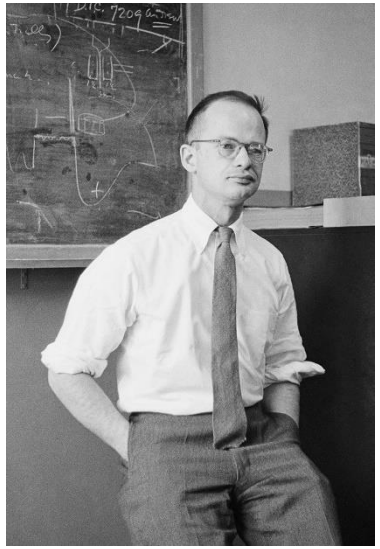
Bulletin of Mathematical Biology Vol. 52, No. 1/2, pp. 99–115, 1990.
Printed in Great Britain.

0092–8240/90$3.00 + 0.00
Pergamon Press plc
Society for Mathematical Biology

A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY*

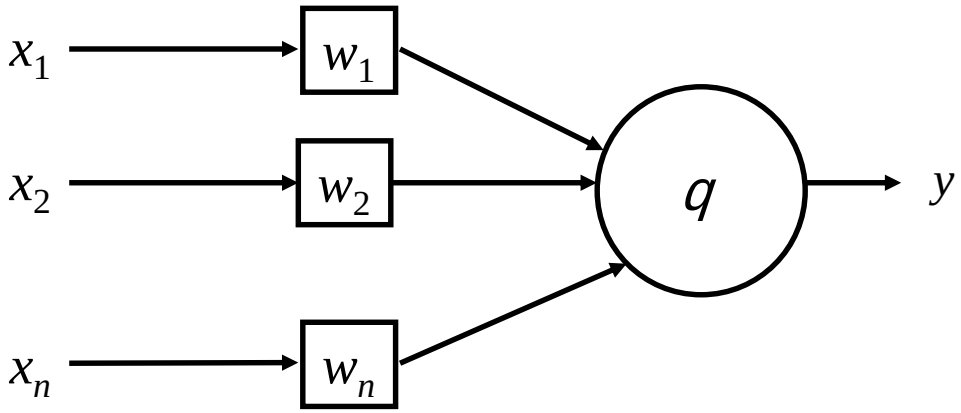
WARREN S. MCCULLOCH AND WALTER PITTS
University of Illinois, College of Medicine,
Department of Psychiatry at the Illinois Neuropsychiatric Institute,
University of Chicago, Chicago, U.S.A.

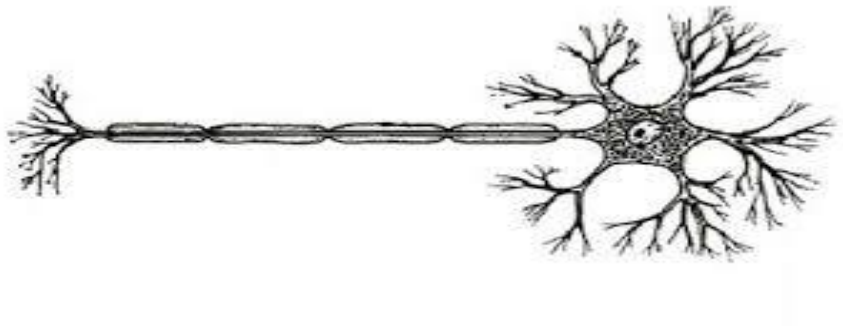




Perceptron

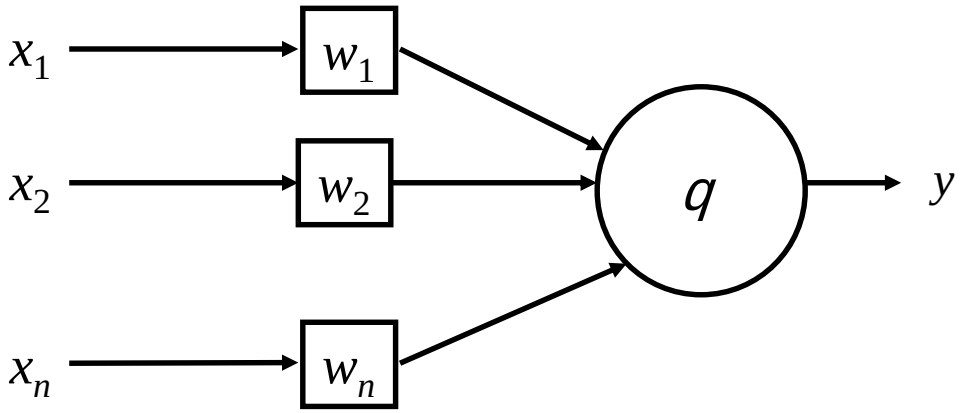
- Modelo de Neurônio (MCP – McCulloch Pitts)



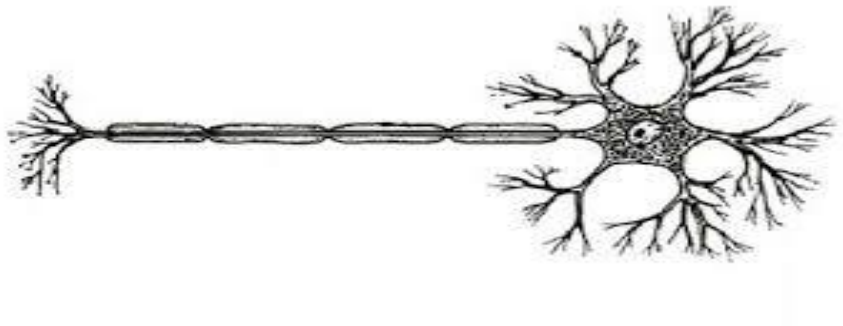


Perceptron

- Modelo de Neurônio (MCP – McCulloch Pitts)

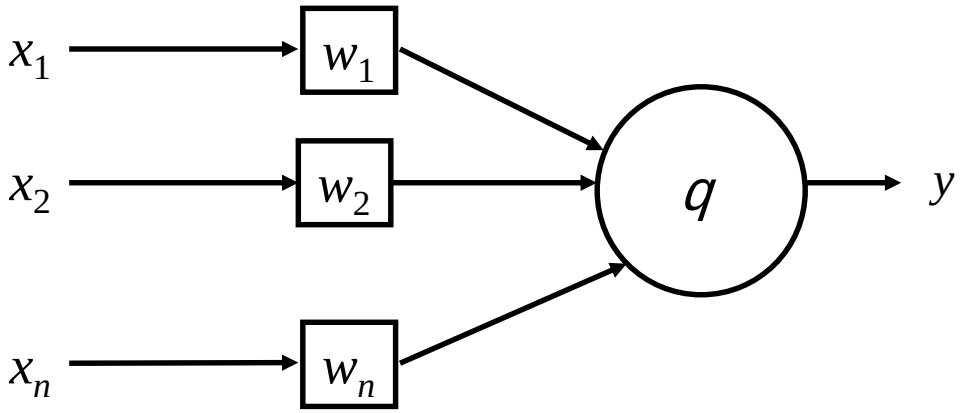


$$y = \theta\left(w_0 + \sum_{i=1}^{d} w_i \cdot x_i\right)$$



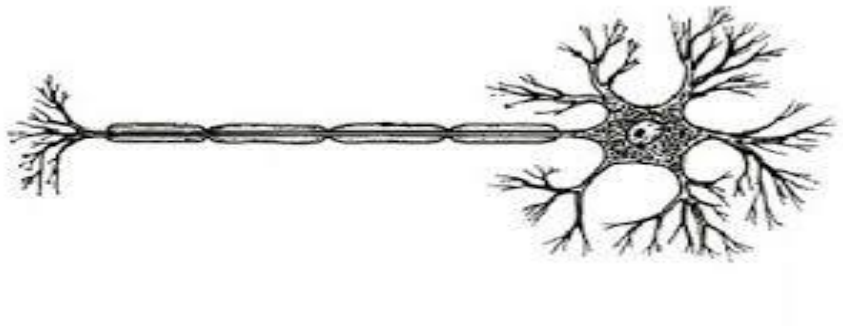
Perceptron

- Modelo de Neurônio (MCP – McCulloch Pitts)



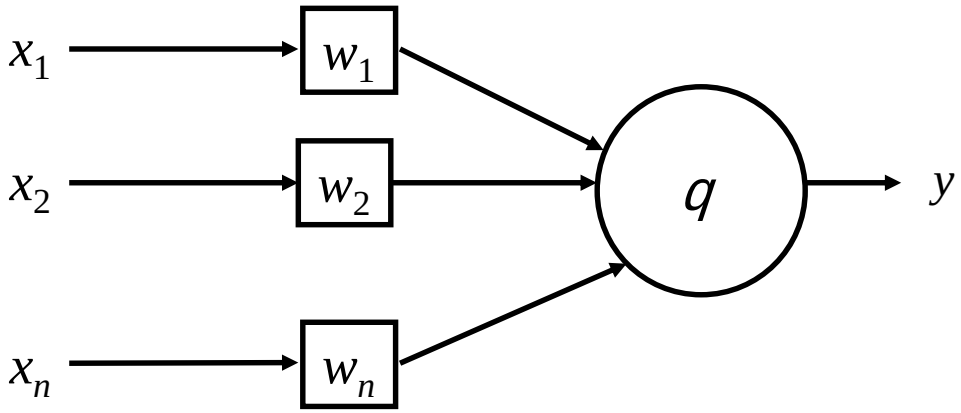
$$y = \theta\left(w_0 + \sum_{i=1}^{d} w_i \cdot x_i\right)$$

$$\theta(x) = \begin{cases} 1 \; se \; x > 0 \\ 0 \quad senão \end{cases}$$



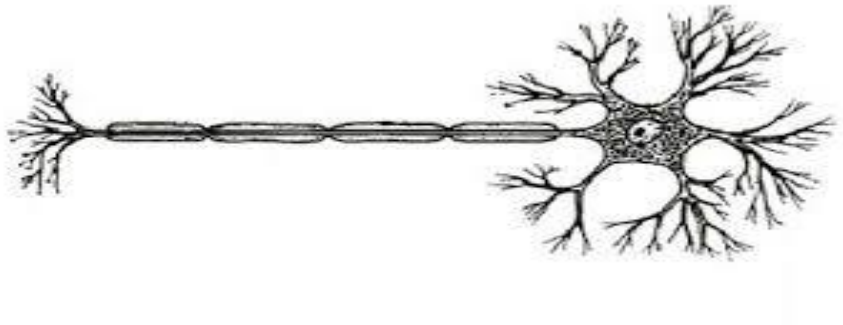
Perceptron

- Modelo de Neurônio (MCP – McCulloch Pitts)



$$y = \theta\left(w_0 + \sum_{i=1}^{d} w_i \cdot x_i\right)$$

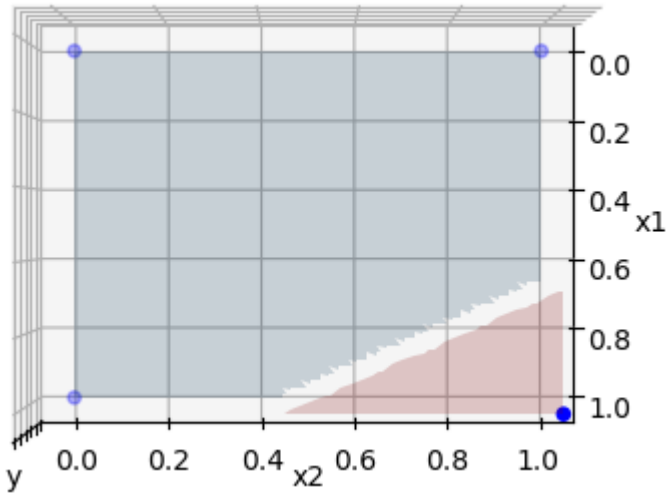
classificador linear!



Perceptron

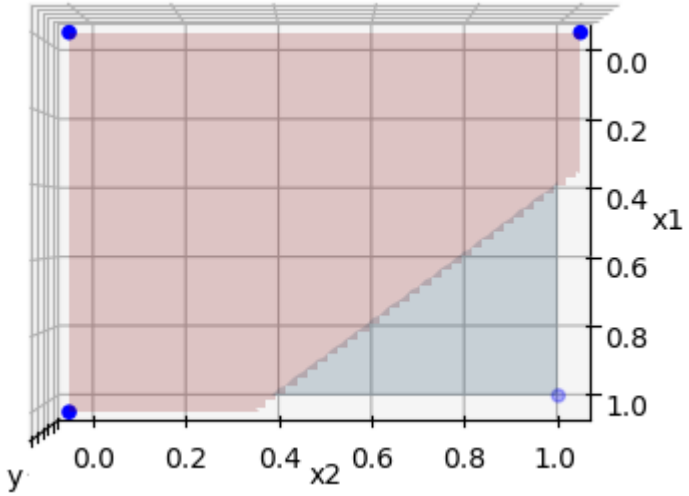
- Modelo de Neurônio (MCP – McCulloch Pitts)
  - Limitação
    - Classifica apenas informações separáveis linearmente

AND



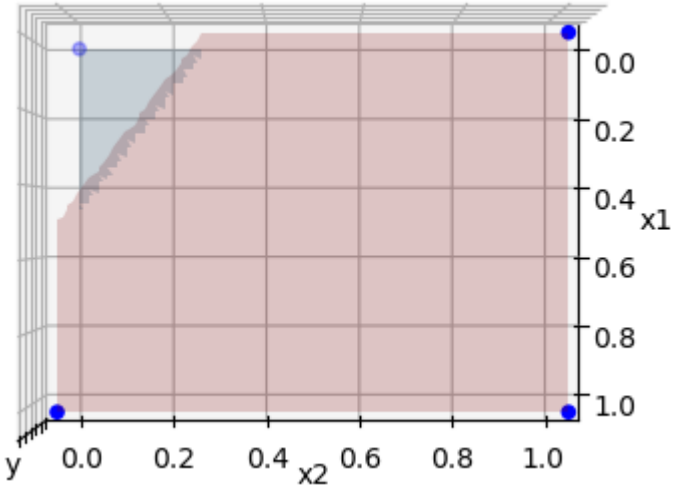
w = [0.70271209 0.422434]
b = [-0.89096912]

OR



w = [1.8462093 3.20945227]
b = [-0.90263086]

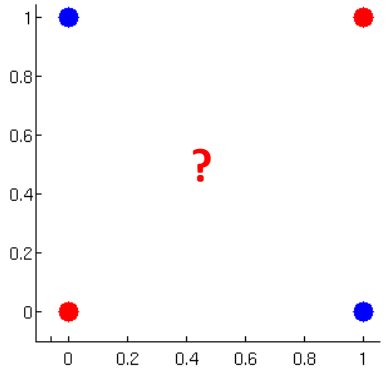
NAND



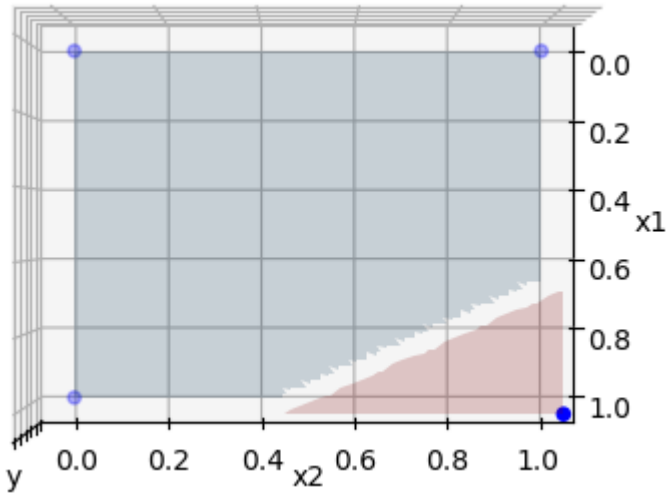
w = [-1.15938768 -1.14587725]
b = [1.5899148]

Perceptron

- Modelo de Neurônio (MCP – McCulloch Pitts)
  - Limitação
    - XOR?

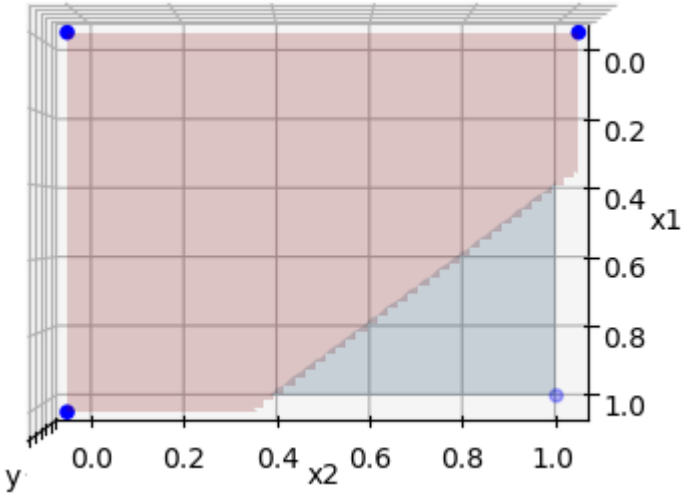


AND



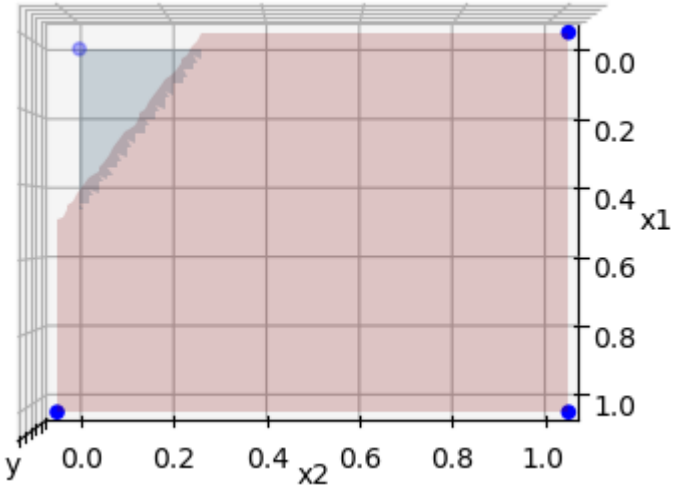
w = [0.70271209 0.422434]
b = [-0.89096912]

OR



w = [1.8462093 3.20945227]
b = [-0.90263086]

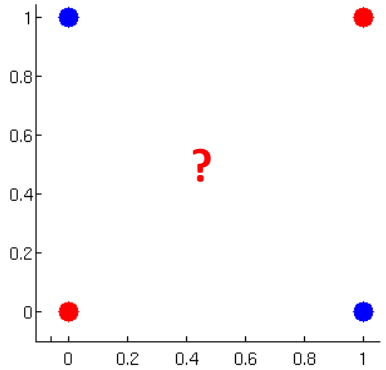
NAND



w = [-1.15938768 -1.14587725]
b = [1.5899148]

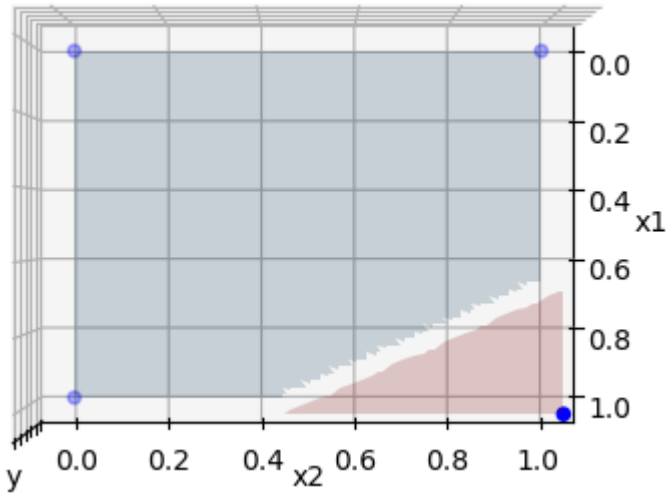
Perceptron

- Modelo de Neurônio (MCP – McCulloch Pitts)
  - Limitação
    - XOR?



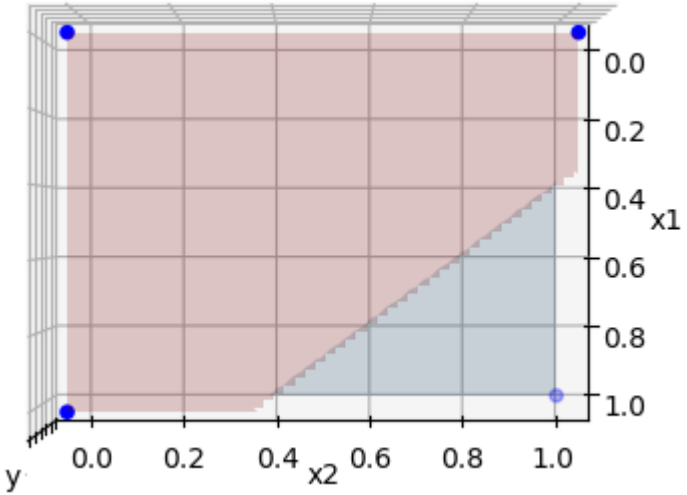
$$p \oplus q \quad = \quad (p \vee q) \wedge \neg(p \wedge q)$$

AND



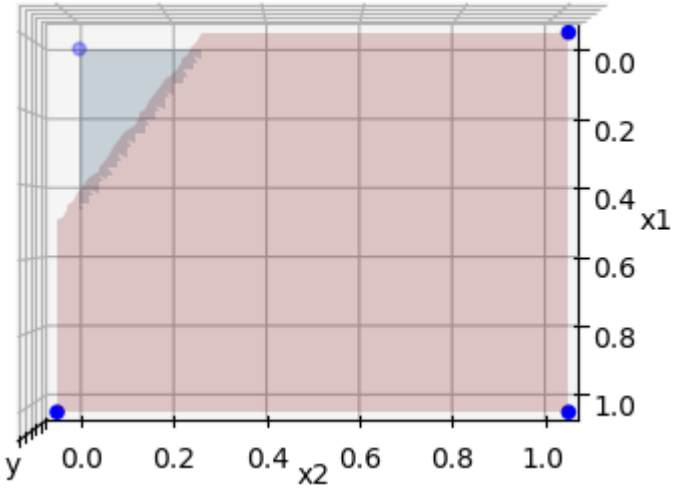
w = [0.70271209 0.422434]
b = [-0.89096912]

OR



w = [1.8462093 3.20945227]
b = [-0.90263086]

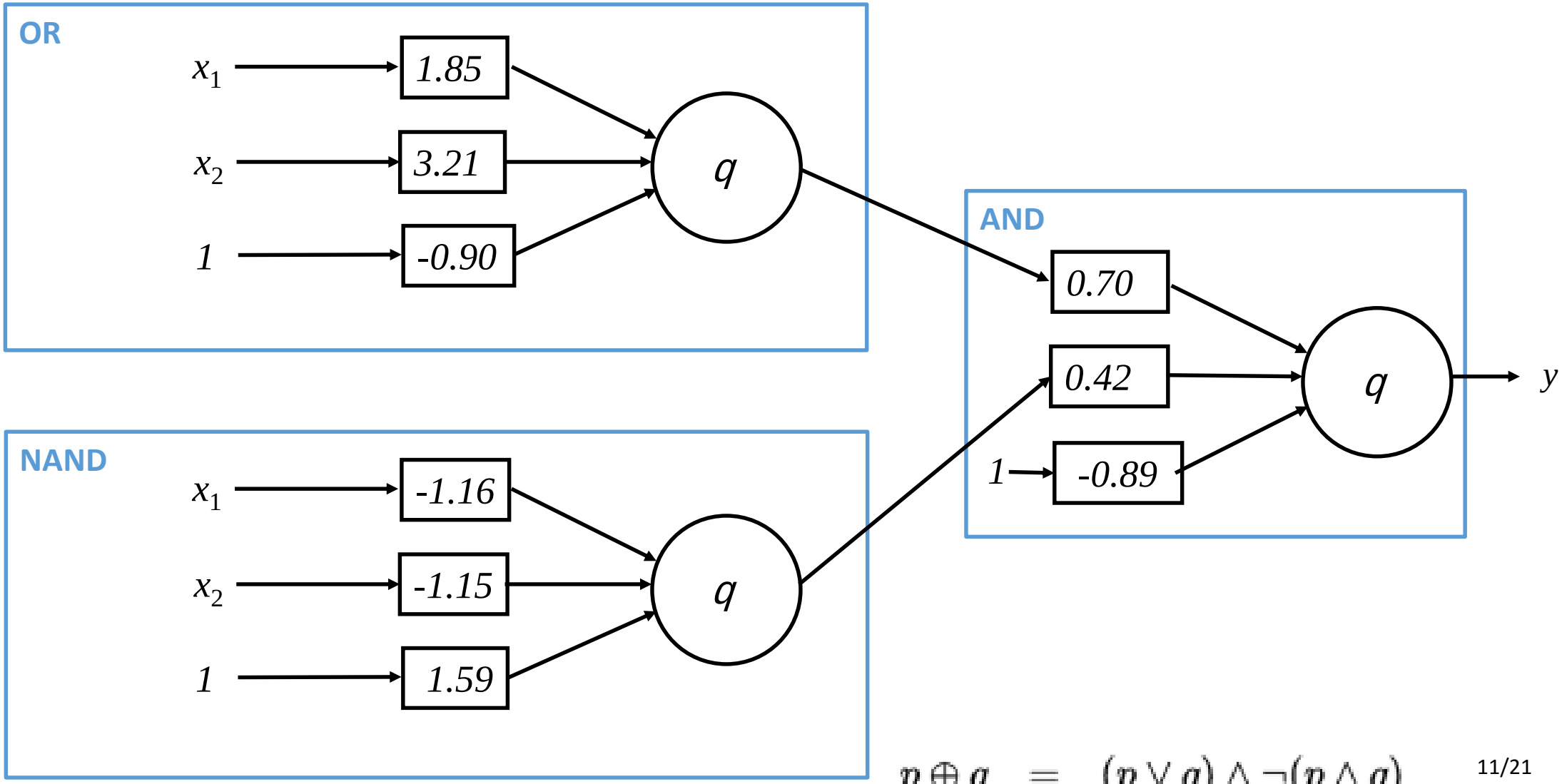
NAND



w = [-1.15938768 -1.14587725]
b = [1.5899148]

Perceptron

- Solução XOR

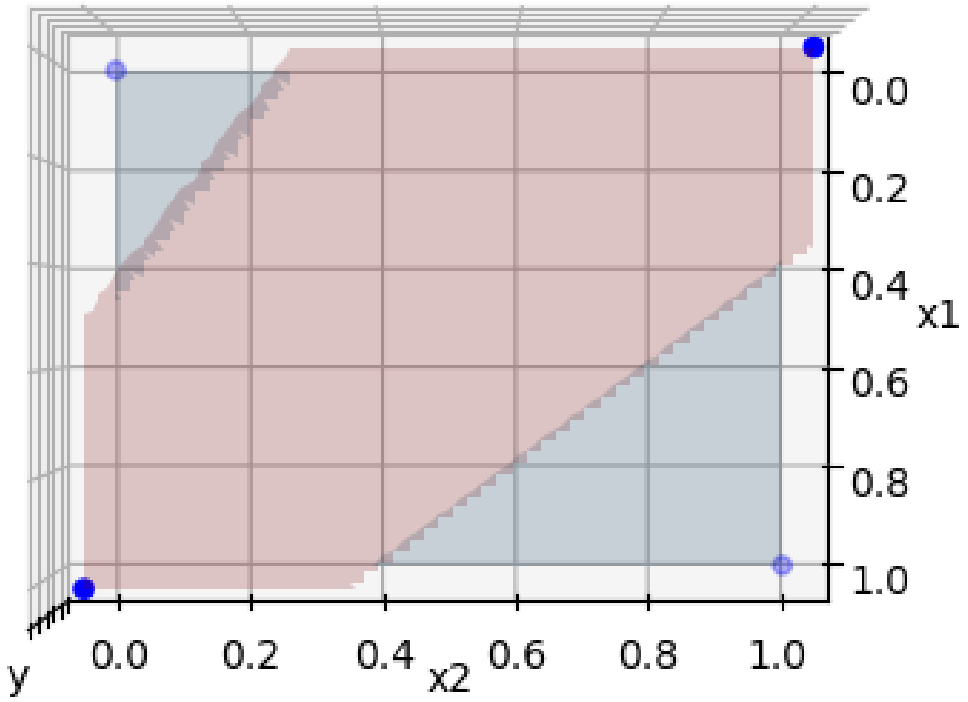


$$p \oplus q \quad = \quad (p \vee q) \wedge \neg(p \wedge q)$$

Perceptron

- Solução XOR

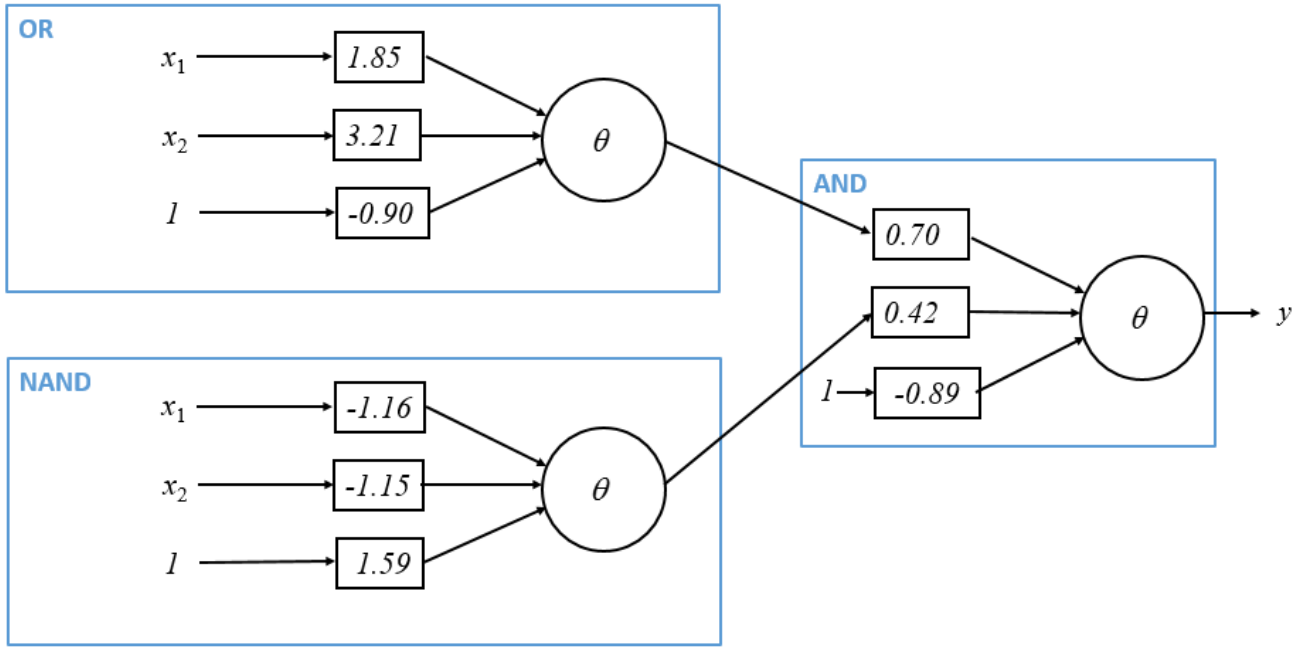
Fronteira de decisão

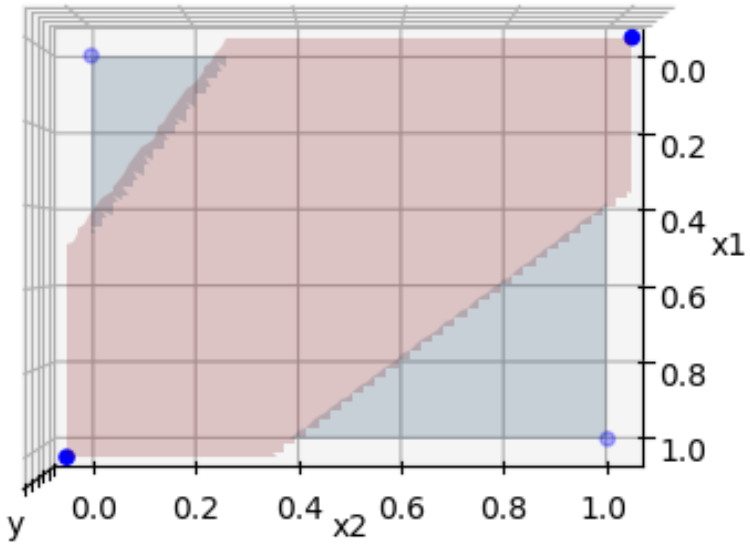


$$p \oplus q \quad = \quad (p \vee q) \wedge \neg(p \wedge q)$$

Perceptron

- Solução XOR

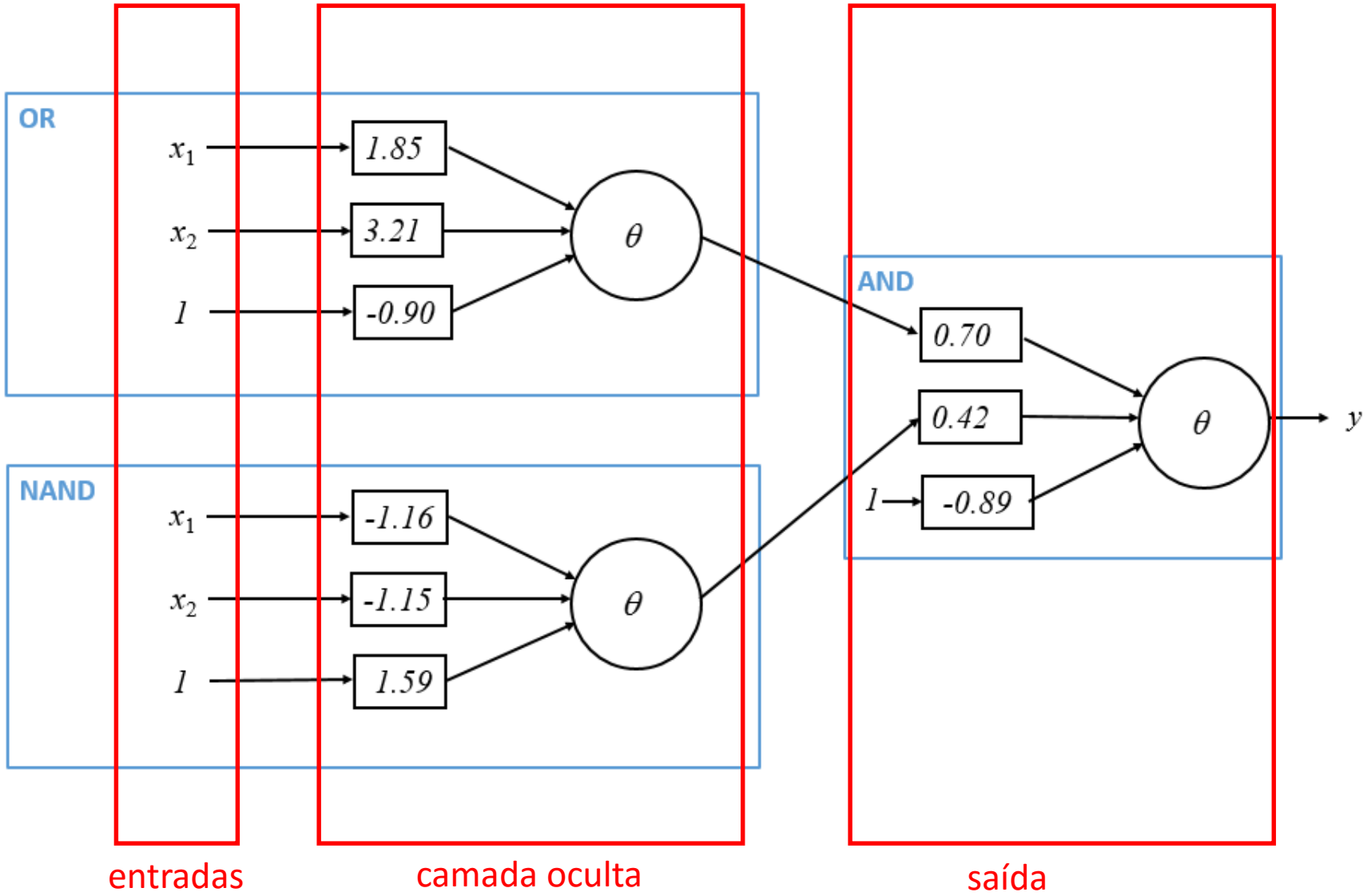




- Duas interpretações
  - Perceptrons mapeiam subproblemas
  - Perceptons definem fronteiras de decisão

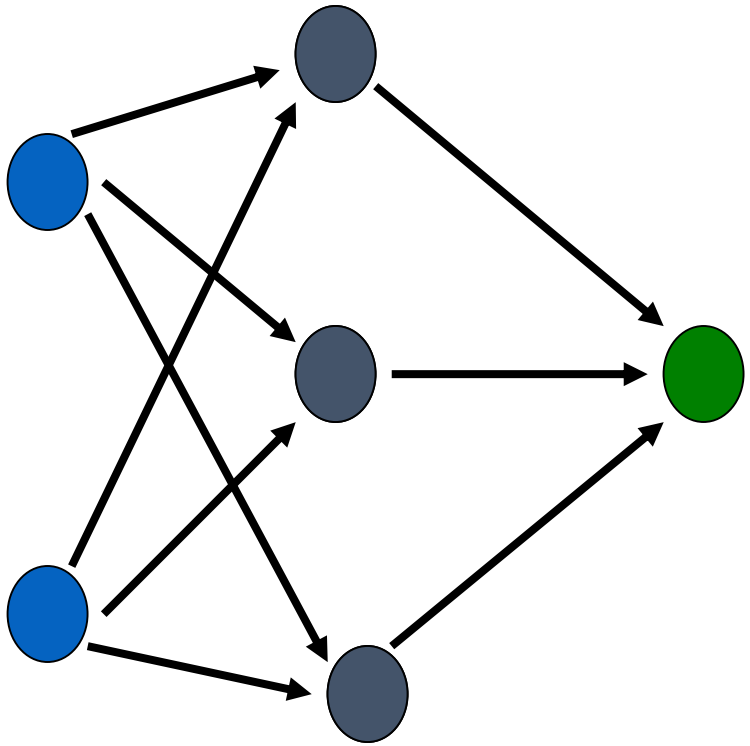
Multilayer Perceptron

- Perceptrons ligados em camadas



Multilayer Perceptron

- Perceptrons ligados em camadas



camada de entrada

camada oculta

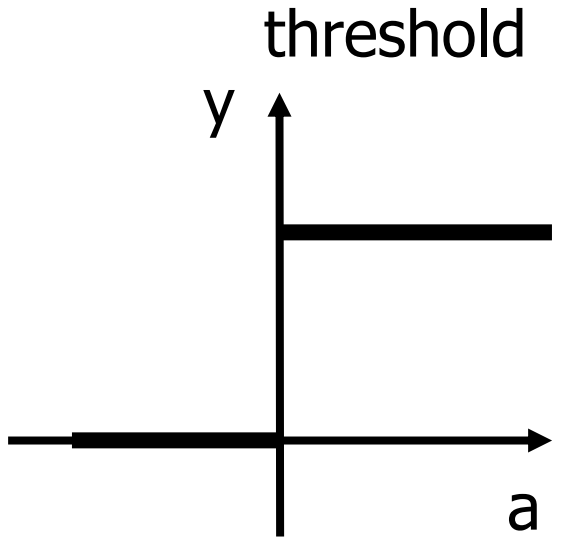
camada de saída

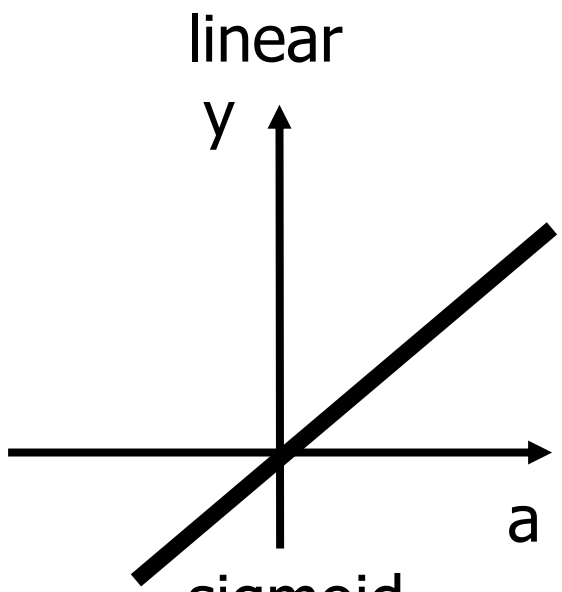
Multilayer Perceptron

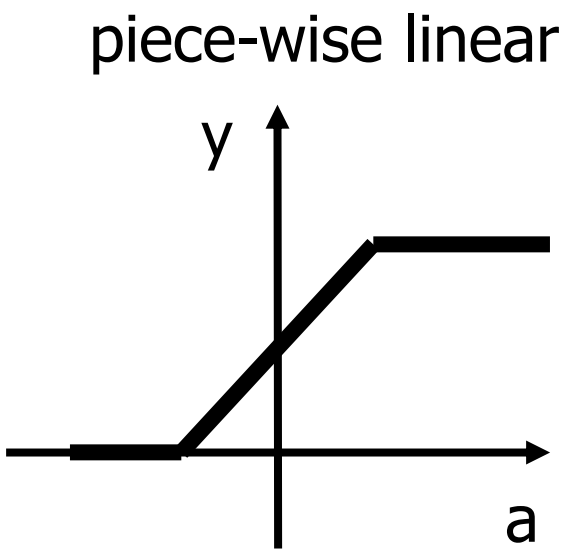
- Perceptrons ligados em camadas
  - Uma camada de entrada
  - Uma camada oculta
  - Uma camada de saída
  - Perceptrons no MLP não são limitados à uma única função de ativação (Heaviside)

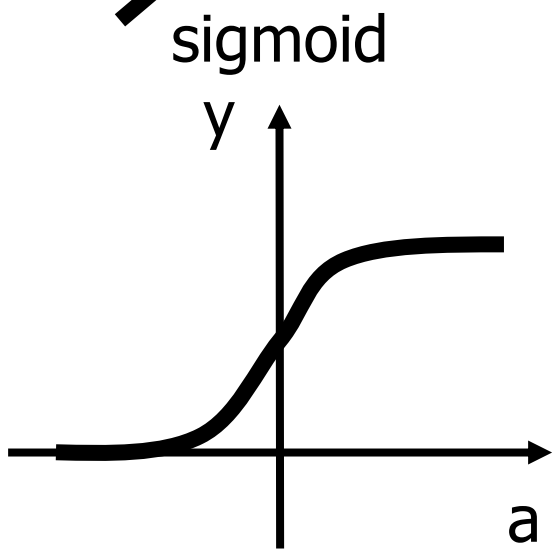
Multilayer Perceptron

- Funções de ativação









Multilayer Perceptron

- Teorema da Aproximação Universal de Funções
  - Teorema da aproximação de Kolmogorov

- Kolmogorov demonstrou que qualquer função contínua g(x) definida no hipercubo unitário In pode ser representada pela escolha adequada de funções Ξ_j e ψ_{ij} por:

$$g(x) = \sum_{j=1}^{2n+1} \Xi_j \left(\sum_{i=1}^{d} \psi_{ij}(x_i) \right)$$

Kolmogorov, A. N. (1957). "On the Representation of Continuous Functions of Several Variables by Superposition of Continuous Functions of one Variable and Addition," Doklady Akademii. Nauk USSR, 114, 679-681.

Multilayer Perceptron

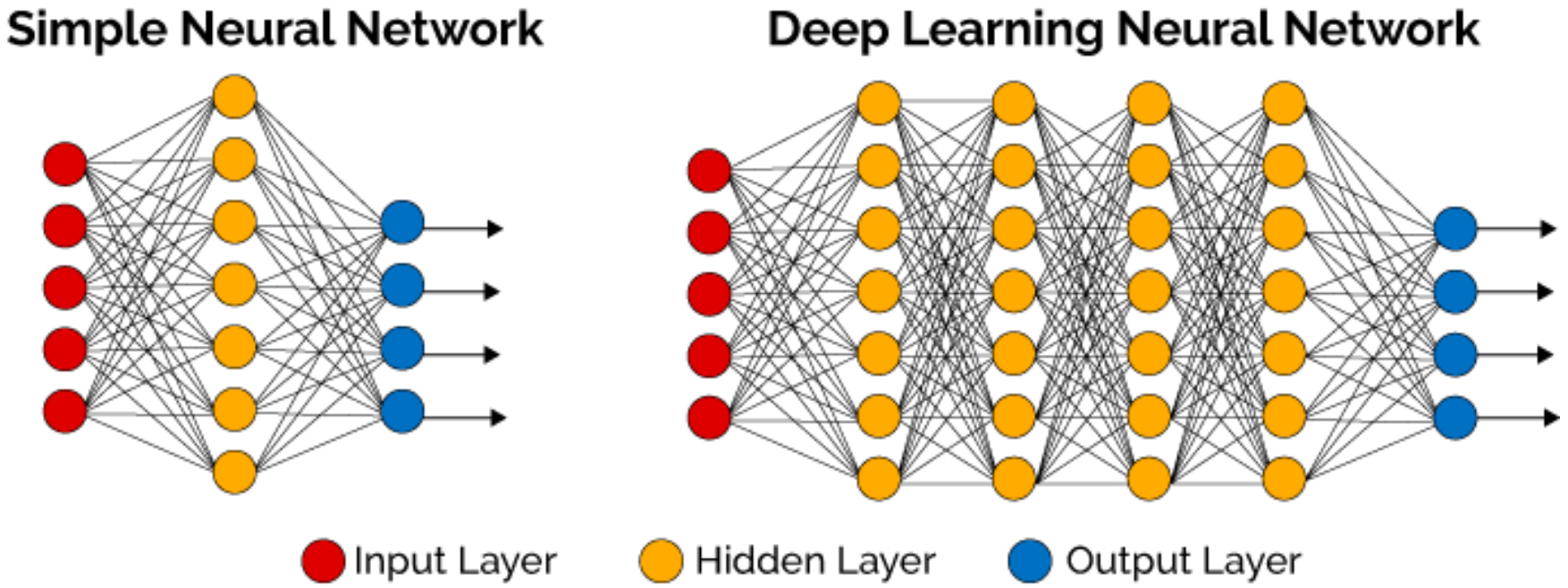
- Teorema da Aproximação Universal de Funções
  - Toda função pode ser aproximada com acurácia arbitrária por uma rede com duas camadas ocultas [Cybenko 1988]
  - Toda função limitada pode ser aproximada com um erro arbitrariamente pequeno, por uma rede com uma camada oculta [Cybenko 1989, Hornik 1989]

Cybenko (1988) Continuous valued neural networks with two hidden layers are sufficient. Technical report. Department of Computer Sciences, Tufts University, Medford, Mass

Cybenko G (1989) Approximation by superposition of a sigmoidal function. Math Control Syst Signals

Multilayer Perceptron

- Deep Learning



Atividade

- Modificar a atividade anterior Breast Cancer

- Utilizar uma framework (Keras, Theano, TensorFlow, PyTorch, etc)