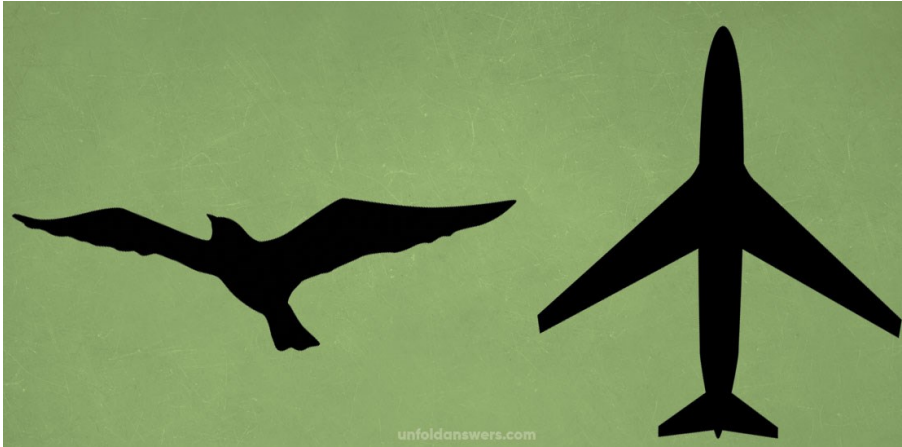


Aprendizado Profundo I

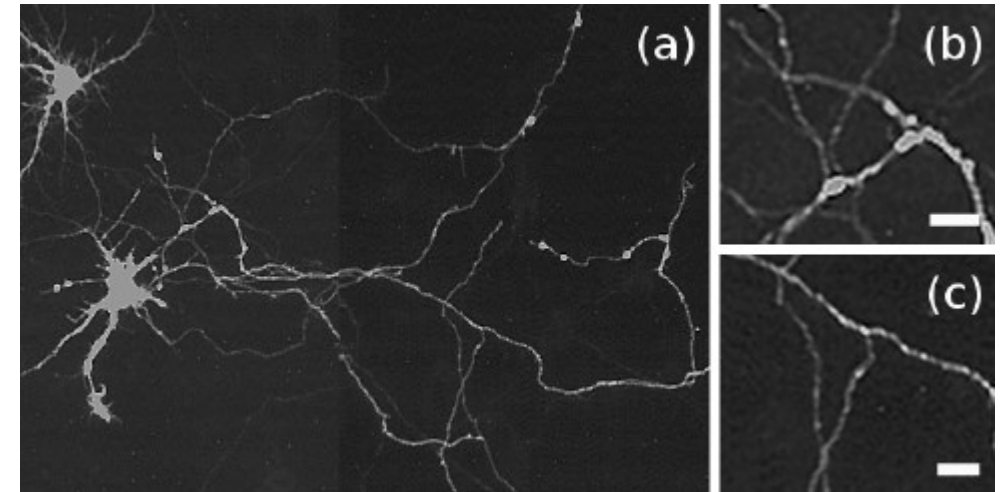
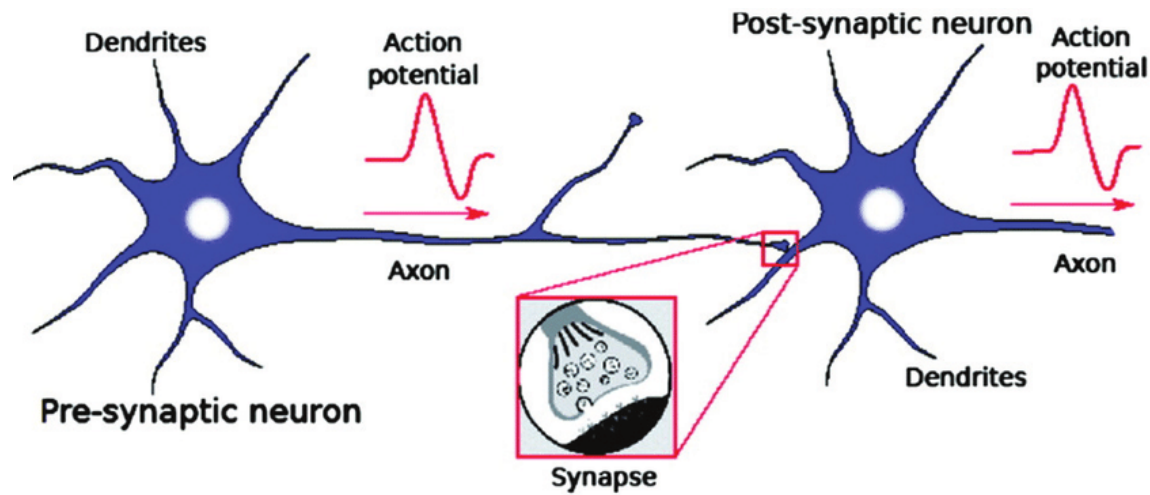
Modelos computacionais do Neurônio: Neurônio MP e Perceptron

Professor: Lucas Silveira Kupssinskü

Motivação - Bioinspirado



Os melhores exemplos de inteligência estão na natureza

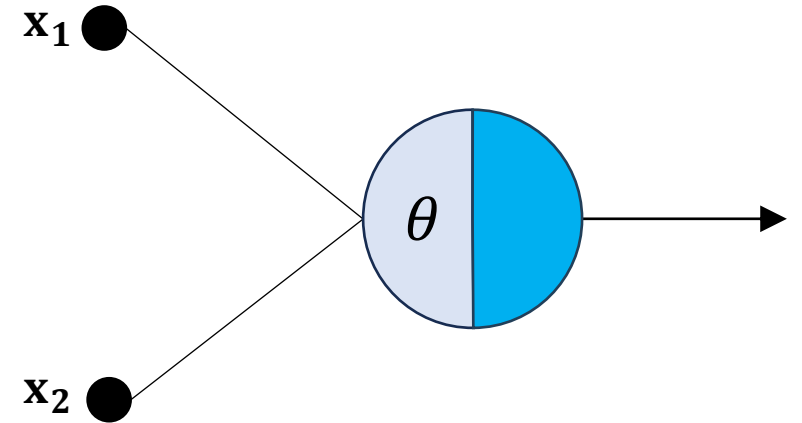


A primeira tentativa – *MP Neuron*

- Importância histórica por ser o primeiro modelo computacional de um neurônio
- Possibilitou mapear a lógica proposicional em neurônios artificiais
- Possível origem do termo “neurônio artificial” e “redes neurais artificiais”
- Modelo teórico, sem implementação prática

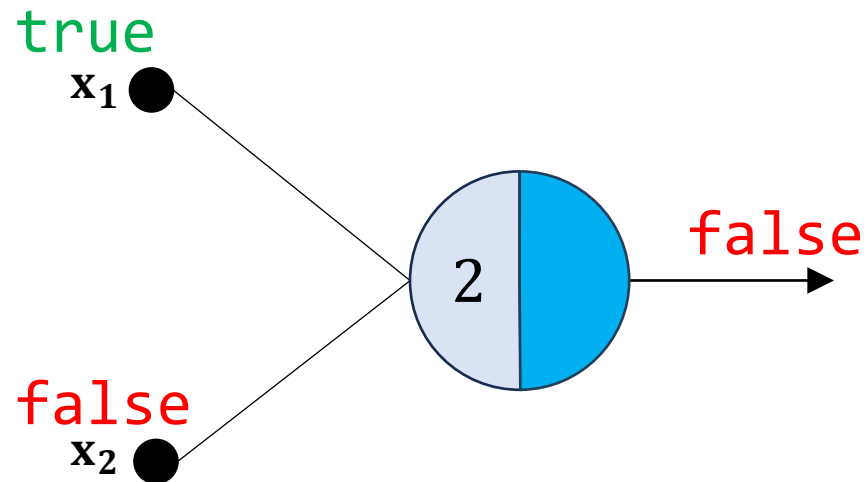
A primeira tentativa – *MP Neuron*

- Entradas binárias
- Estímulos podem ser:
 - Excitatórios: Verdadeiros
 - Inibitórios : Falsos
- Neurônio dispara quando a quantidade de estímulos excitatórios ultrapassar determinado limiar θ



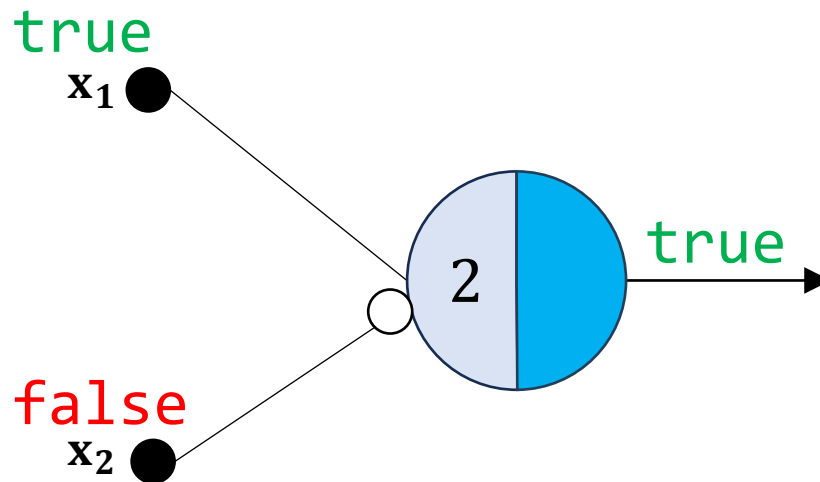
A primeira tentativa – *MP Neuron*

- Exemplo
 - Com limiar $\theta = 2$ o *neurônio MP* vira uma porta lógica *and*



A primeira tentativa – *MP Neuron*

- Exemplo
 - Com uma entrada inibitória
 - Equivalente a aplicar a negação em uma das entradas



A ~~atual~~ segunda tentativa – *Perceptron*

- Pequenas modificações no Neurônio MP (alguns autores não diferenciam os modelos)
 - Cada entrada pode ser um número real qualquer
 - Existe um algoritmo de aprendizado
 - O limiar de decisão não é “projetado” é aprendido

“

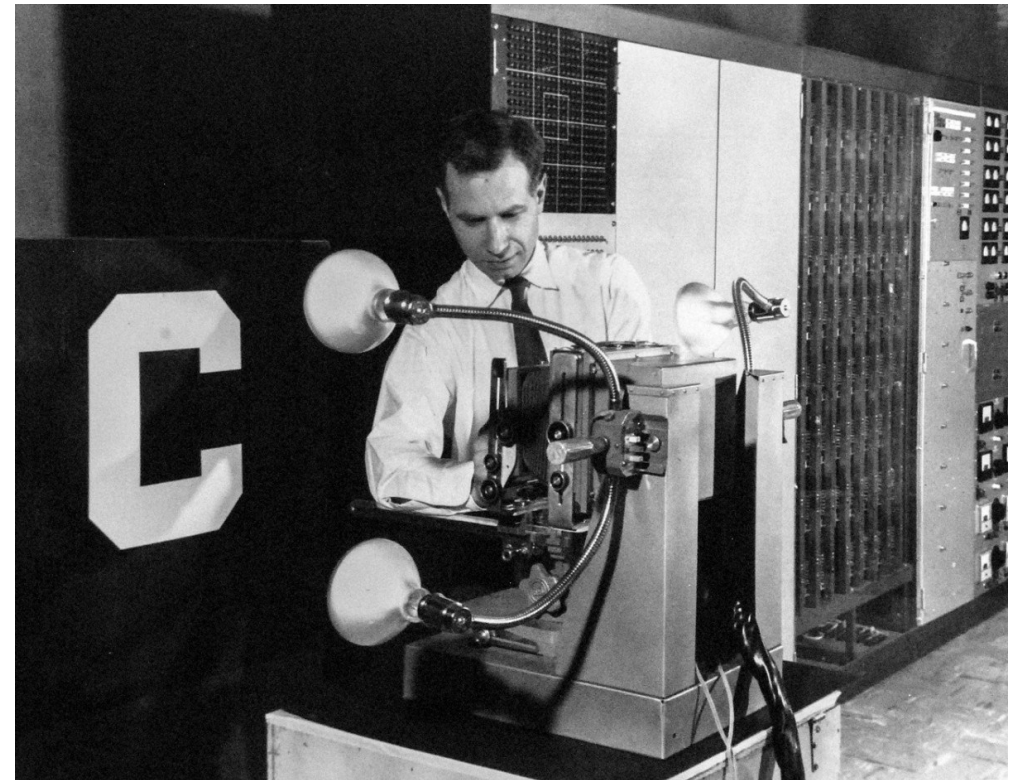
In general, identities of this

”

sort must be learned, or acquired from experience, and if the system is to be economical, the number of functional units in the storage system, or memory, should be much less than the number of forms or memories to be retained.

A segunda tentativa – *Perceptron*

- Foi montado em um *hardware* dedicado para discriminar imagens
 - Resolução 20x20
 - 400 fotocélulas capturavam a imagem
 - Os pesos eram ajustados em potenciômetros



A segunda tentativa – *Perceptron*

- Observação

- Primeira publicação

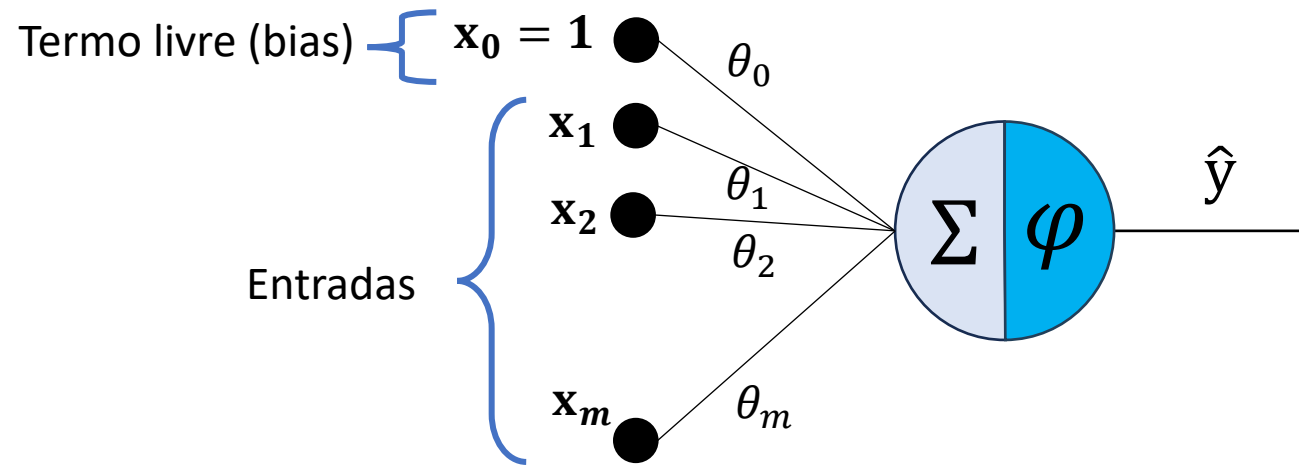
- ROSENBLATT, Frank. **The perceptron, a perceiving and recognizing automaton Project Para**. Cornell Aeronautical Laboratory, 1957.

- Reescrita mais completa da primeira publicação

- ROSENBLATT, Frank. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, v. 65, n. 6, p. 386, 1958.

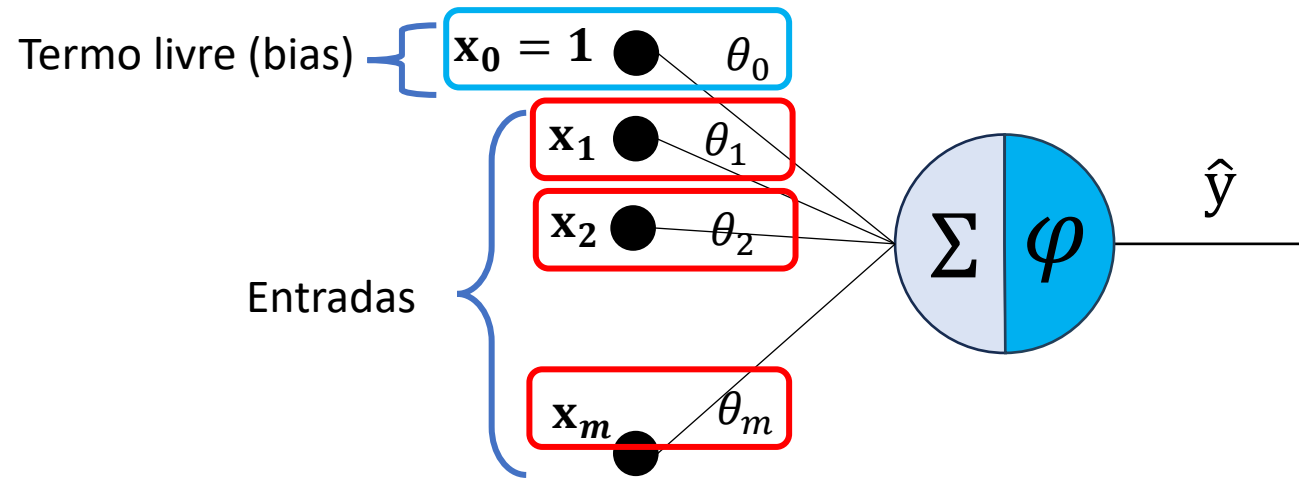
O Perceptron

- $\hat{y}^{(i)} = \varphi(\theta_0 + x_1\theta_1 + x_2\theta_2 + \cdots + x_m\theta_m)$



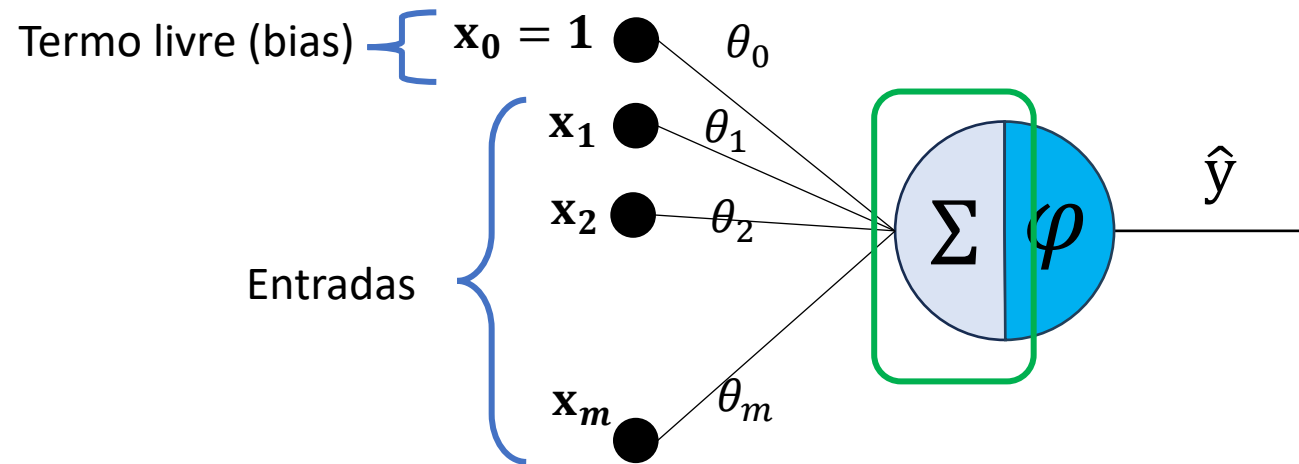
O Perceptron

- $\hat{y}^{(i)} = \varphi(\theta_0 + x_1\theta_1 + x_2\theta_2 + \dots + x_m\theta_m)$



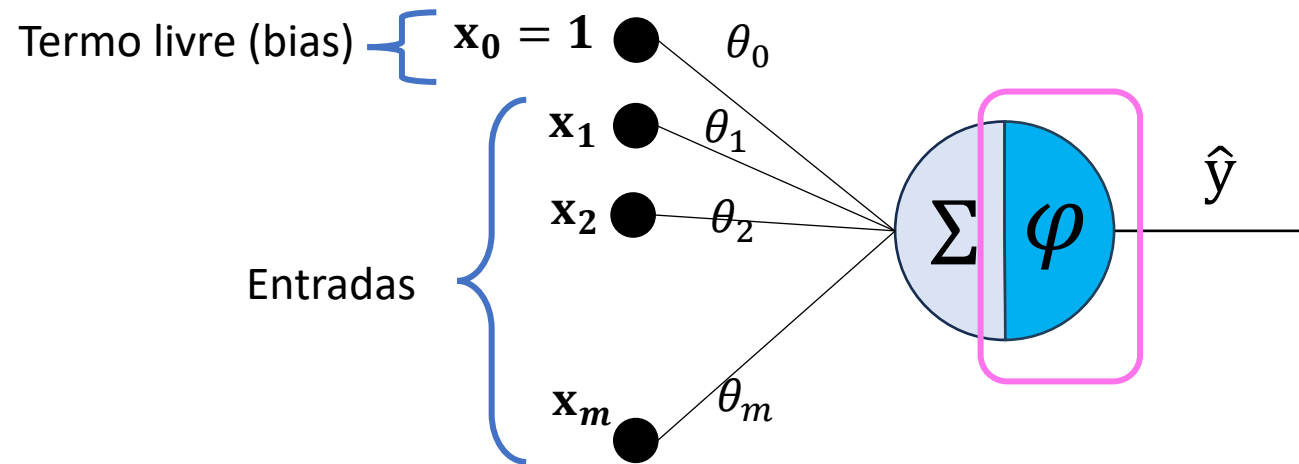
O Perceptron

- $\hat{y}^{(i)} = \varphi(\theta_0 + x_1\theta_1 + x_2\theta_2 + \cdots + x_m\theta_m)$



O Perceptron

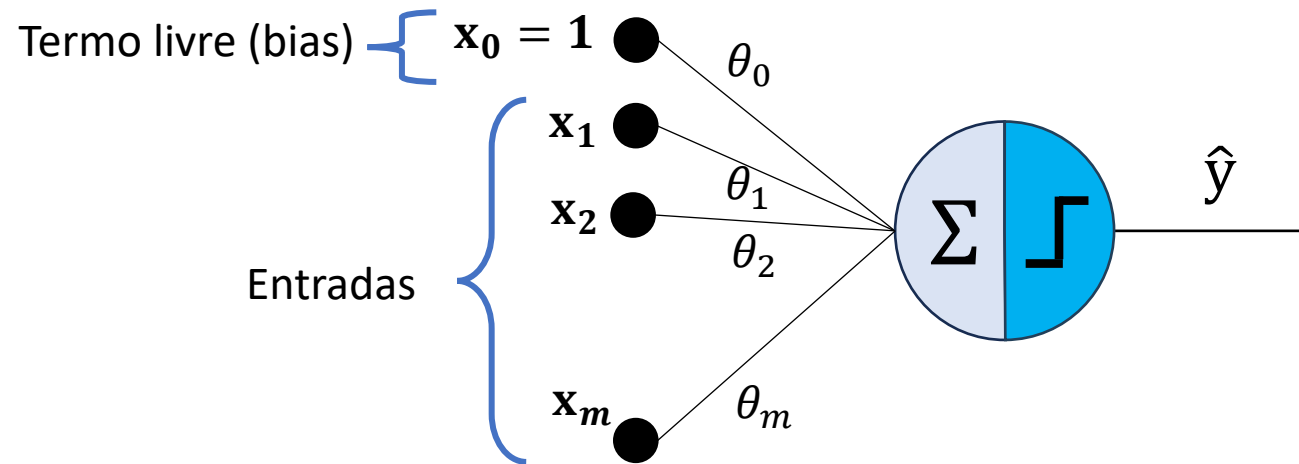
- $\hat{y}^{(i)} = \varphi(\theta_0 + x_1\theta_1 + x_2\theta_2 + \cdots + x_m\theta_m)$



O Perceptron de Rosenblatt

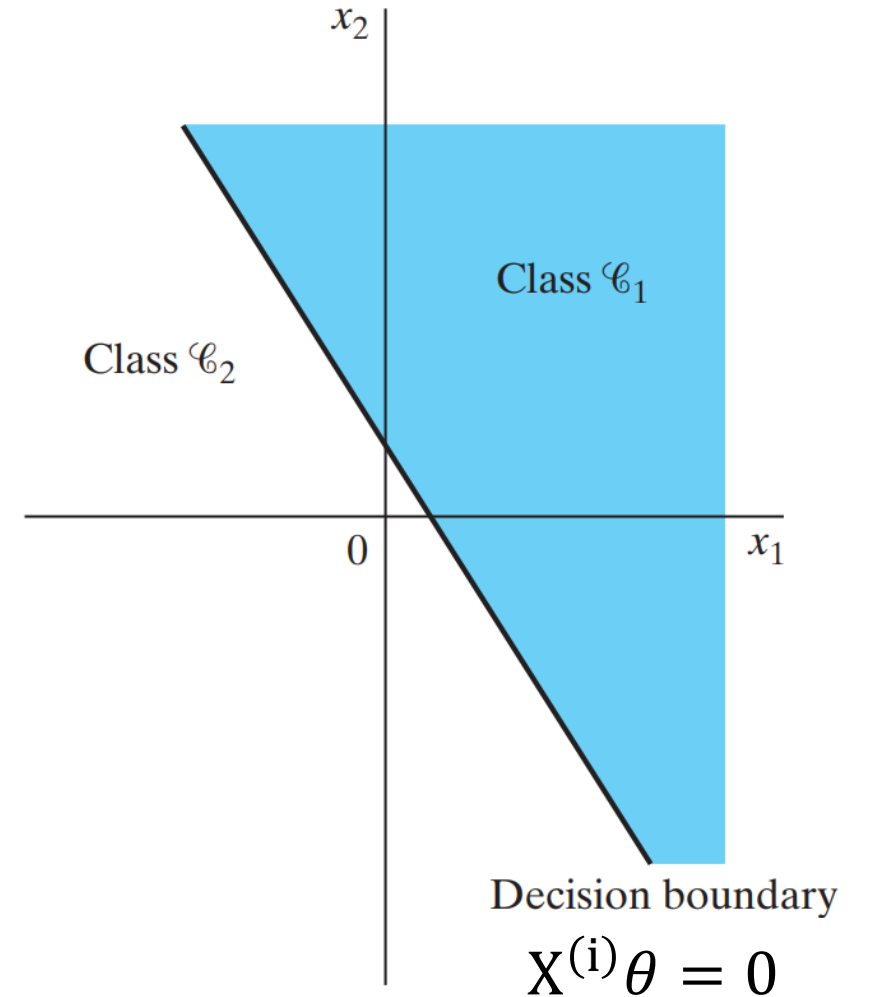
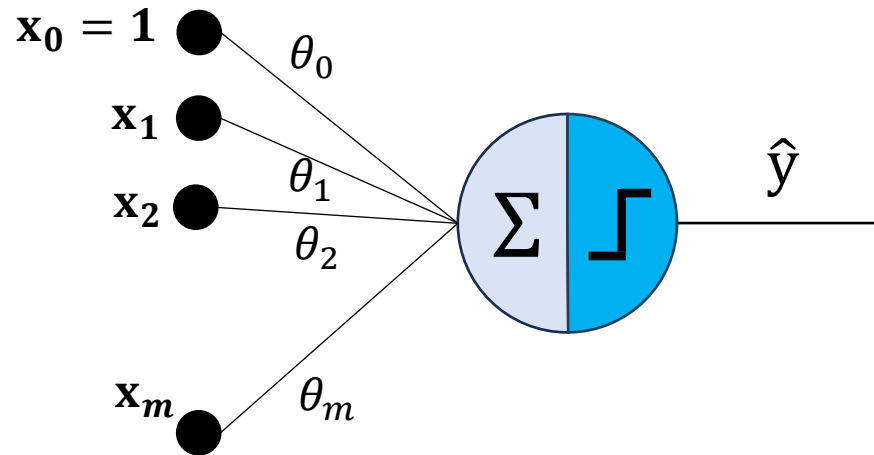
- $\hat{y}^{(i)} = \text{sign}(X^{(i)} \theta)$

$$\text{sign}(x) = \begin{cases} +1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$



O Perceptron de Rosenblatt

- $\hat{y}^{(i)} = \text{sign}(X^{(i)}\theta)$



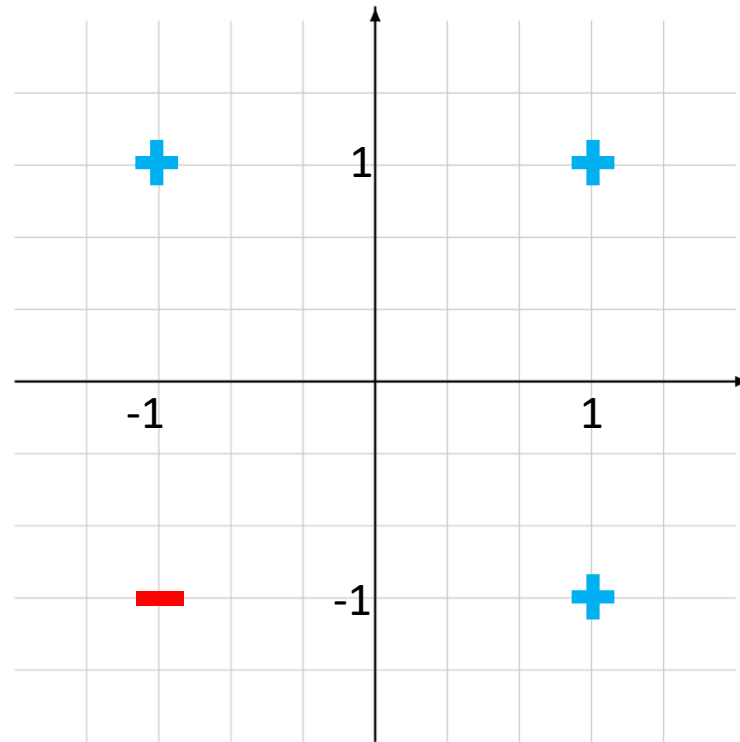
Perceptron Learning Algorithm (PLA)

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$

Esse algoritmo converge desde que os dados sejam linearmente separáveis.

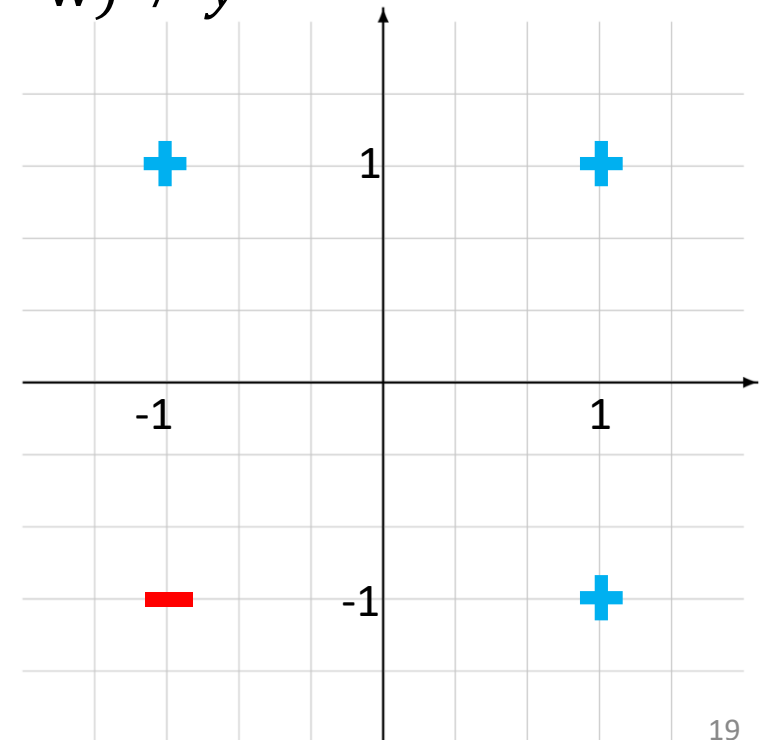
Exemplo

X1	X2	Y
-1	-1	-1
-1	1	1
1	-1	1
1	1	1



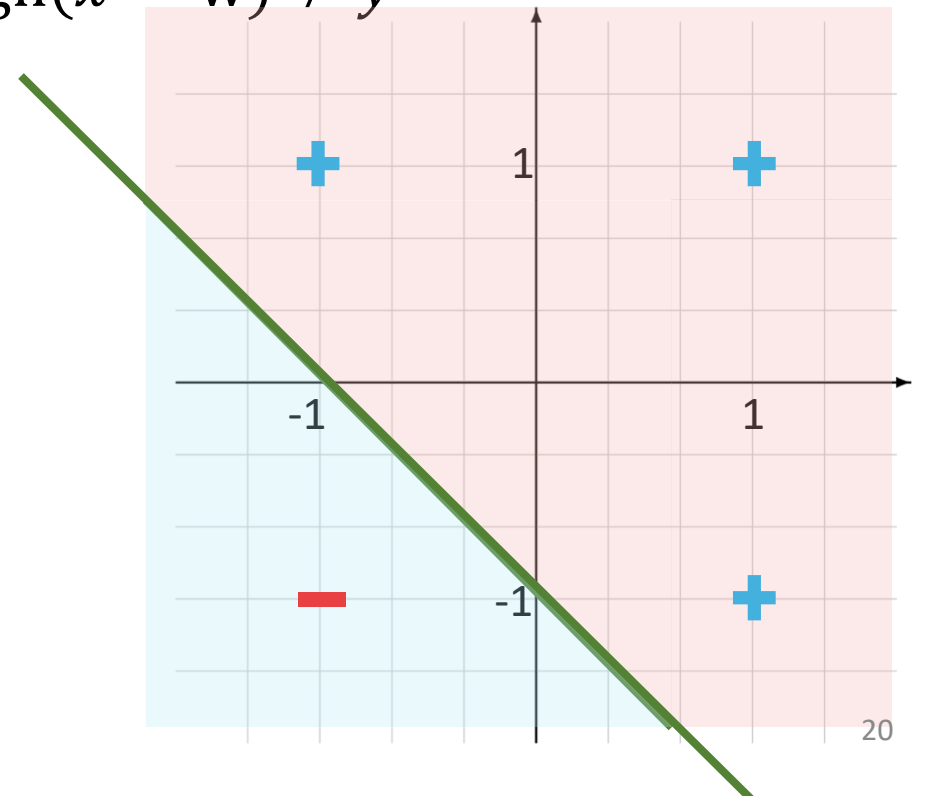
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$



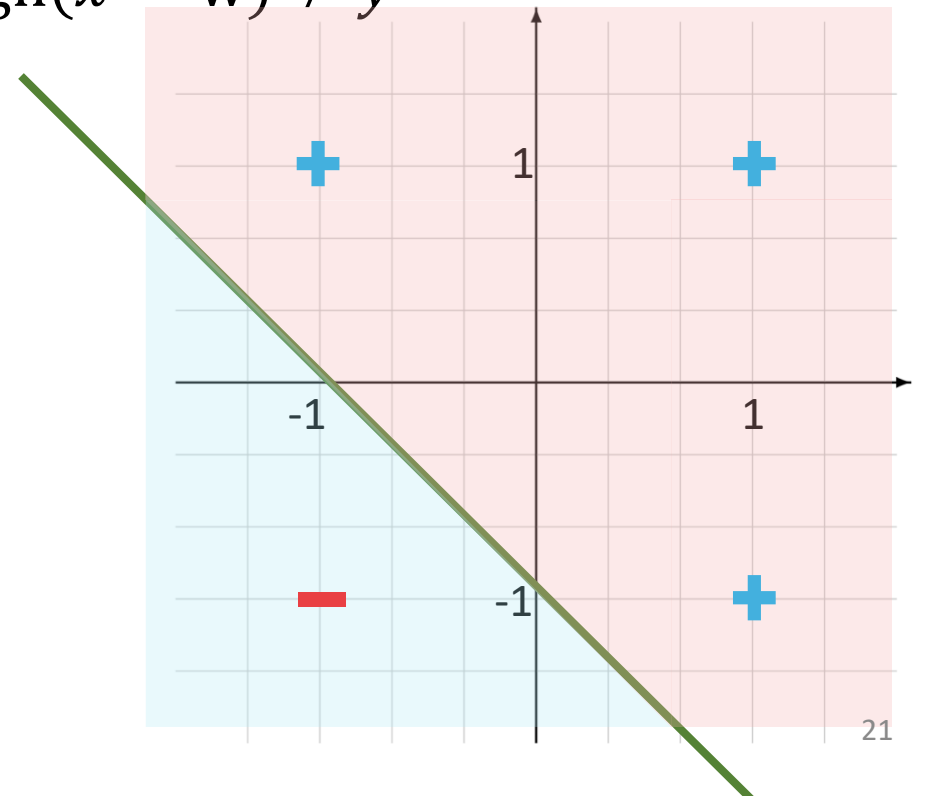
Exemplo

- ▶ Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [-1, -1, -1]$



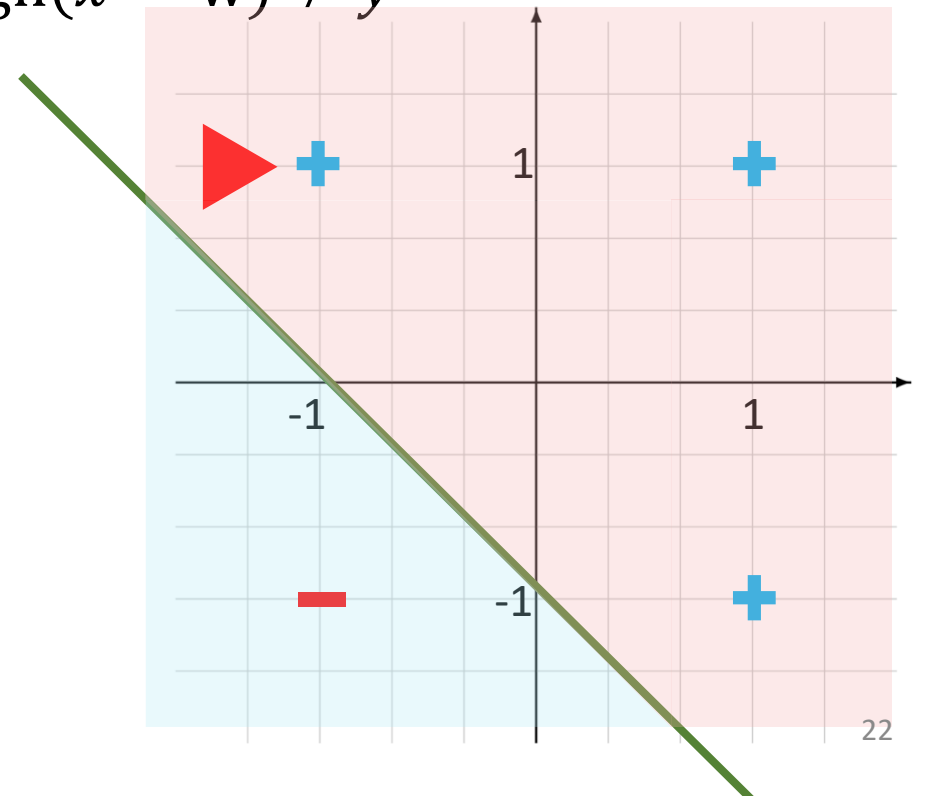
Exemplo

- Inicialize os pesos w aleatoriamente;
- ▶ Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [-1, -1, -1]$



Exemplo

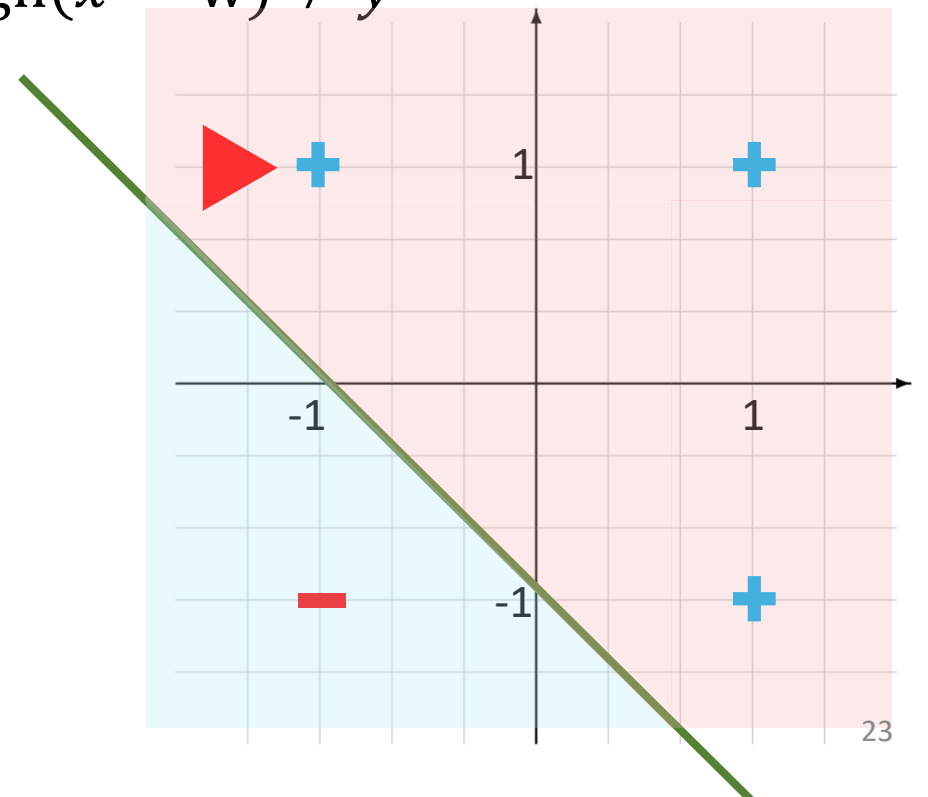
- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - ▶ Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [-1, -1, -1]$



Exemplo

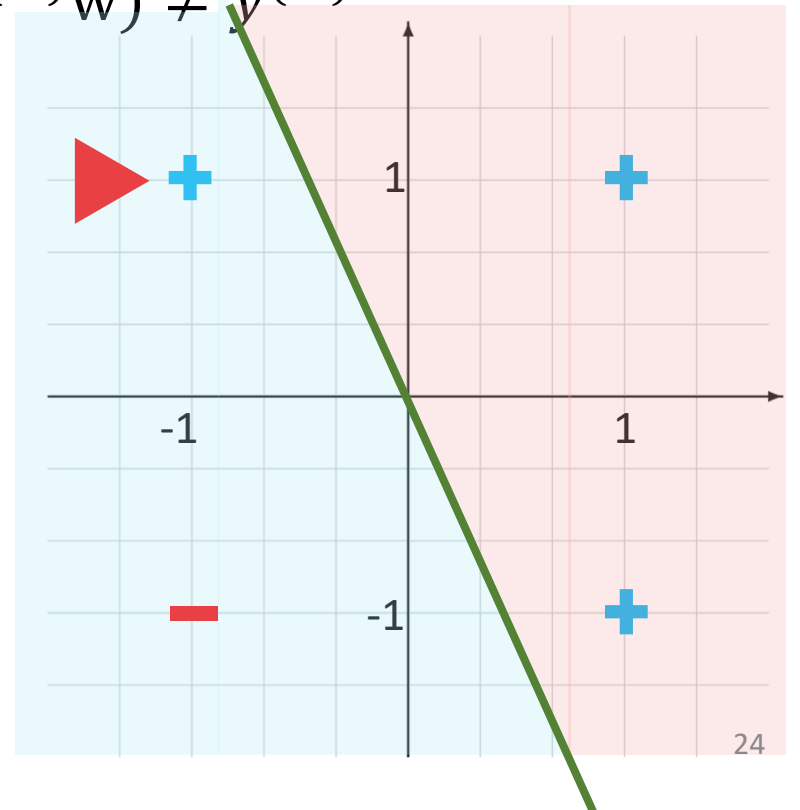
- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - ▶ Atualize os pesos: $w = w + y^{(m)}x^{(m)}$

- $w^T = [-1, -1, -1]$
- $w = [-1, -1, -1] + [1, -1, 1]$



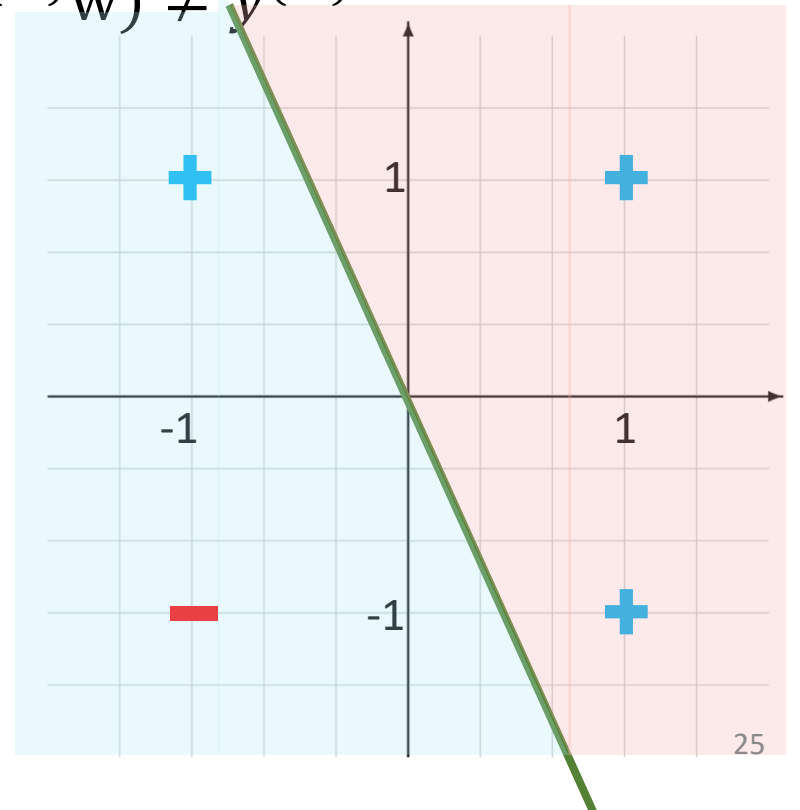
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - ▶ Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [0, -2, 0]$



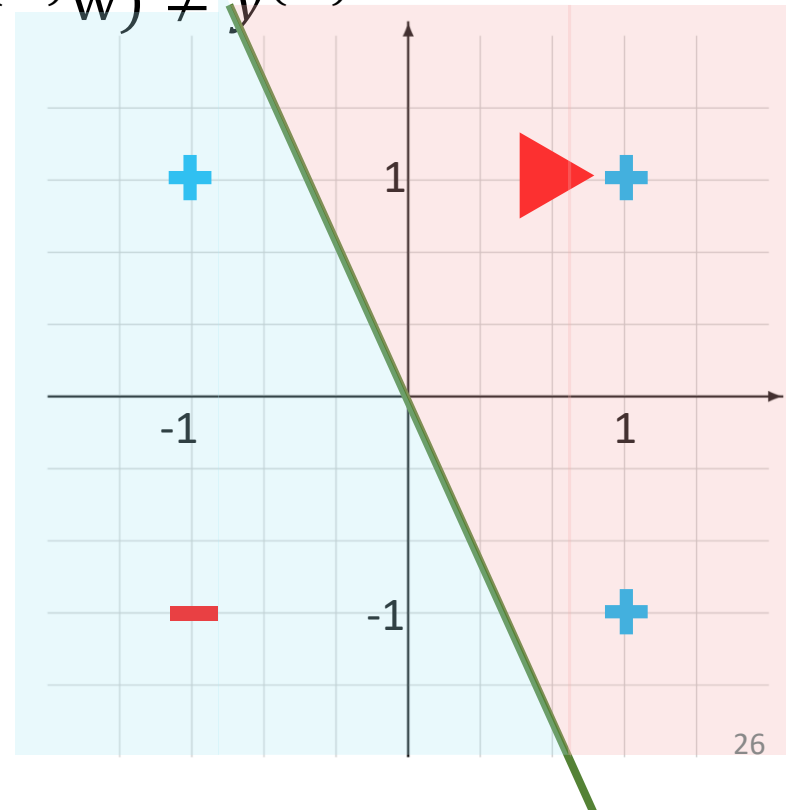
Exemplo

- Inicialize os pesos w aleatoriamente;
- ▶ Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [0, -2, 0]$



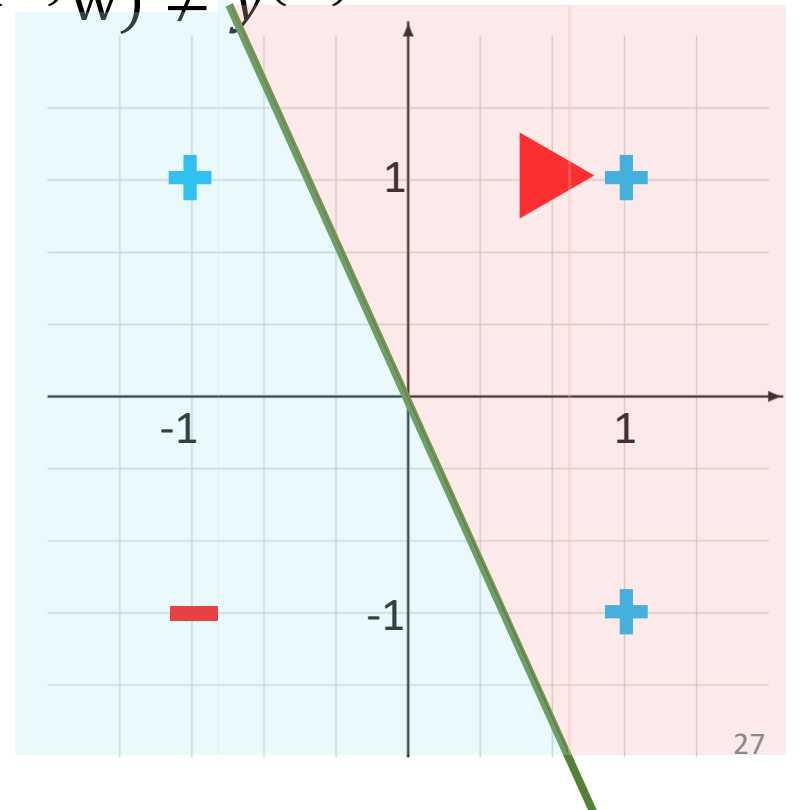
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - ▶ Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [0, -2, 0]$



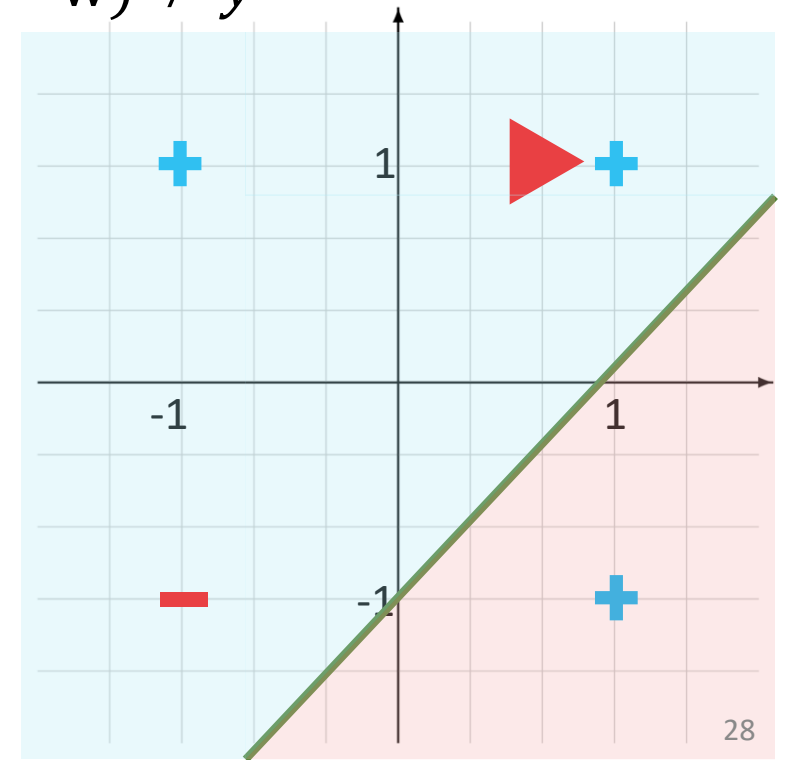
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - ▶ Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [0, -2, 0]$
- $w = [0, -2, 0] + [1, 1, 1]$



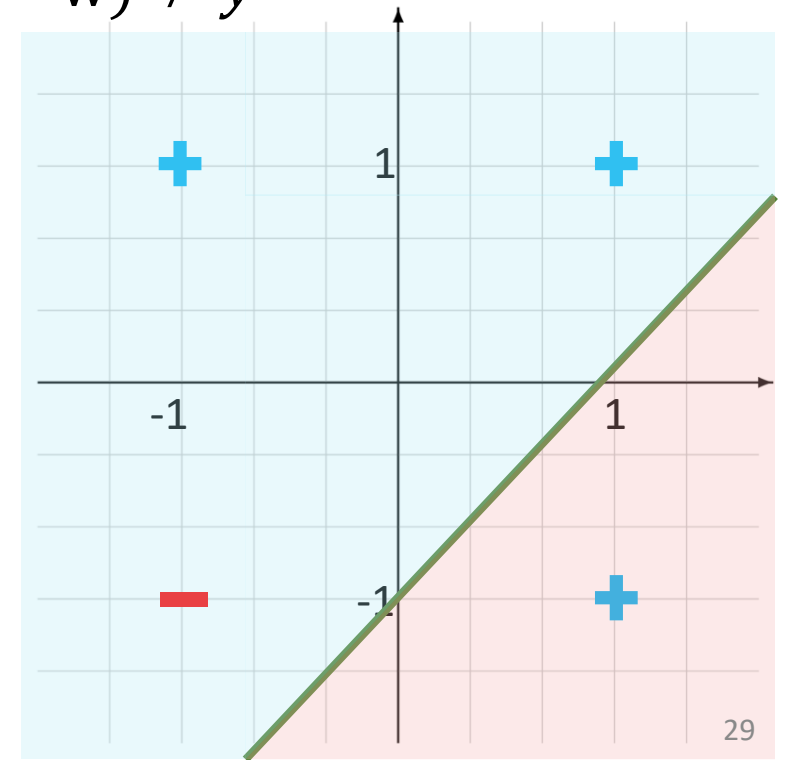
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - ▶ Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [1, -1, 1]$



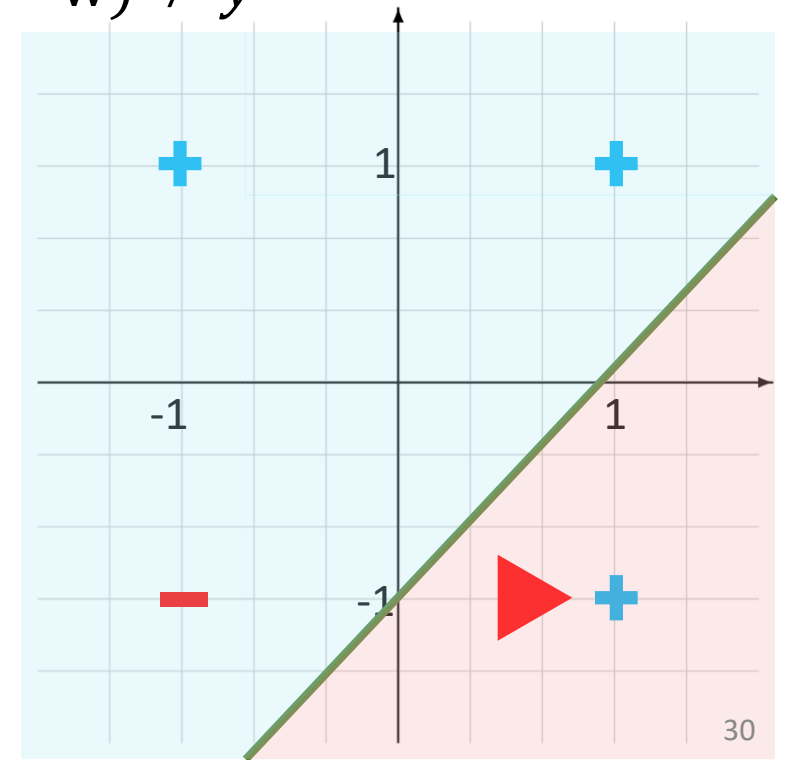
Exemplo

- Inicialize os pesos w aleatoriamente;
- ▶ Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [1, -1, 1]$



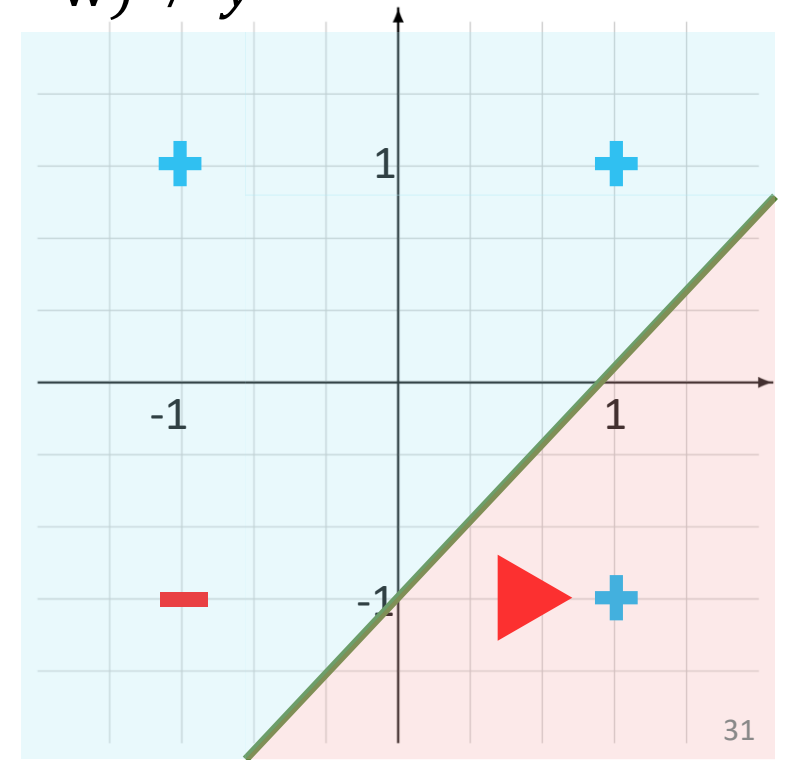
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - ▶ Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [1, -1, 1]$



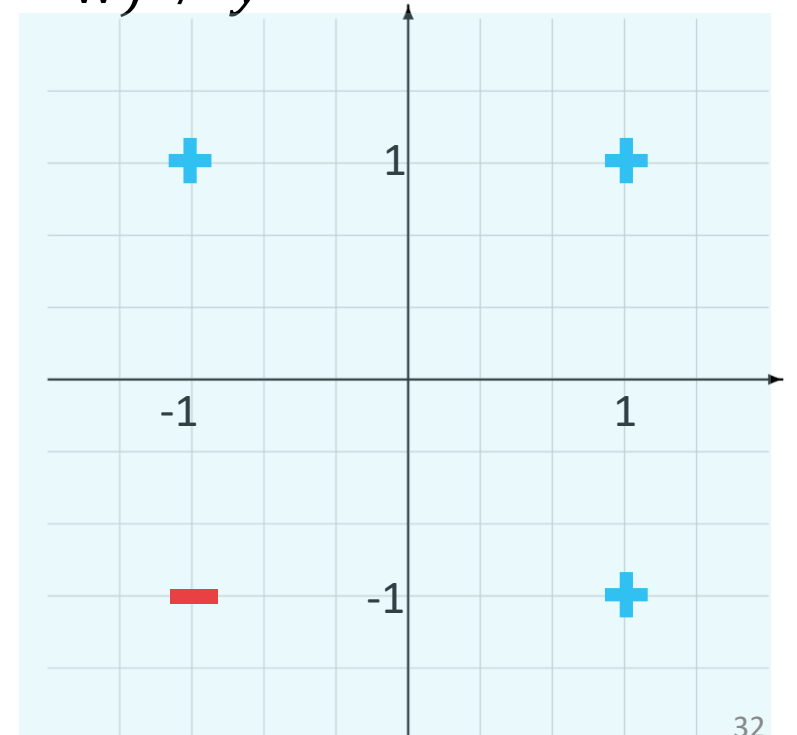
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - ▶ Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [1, -1, 1]$
- $w = [1, -1, 1] + [1, 1, -1]$



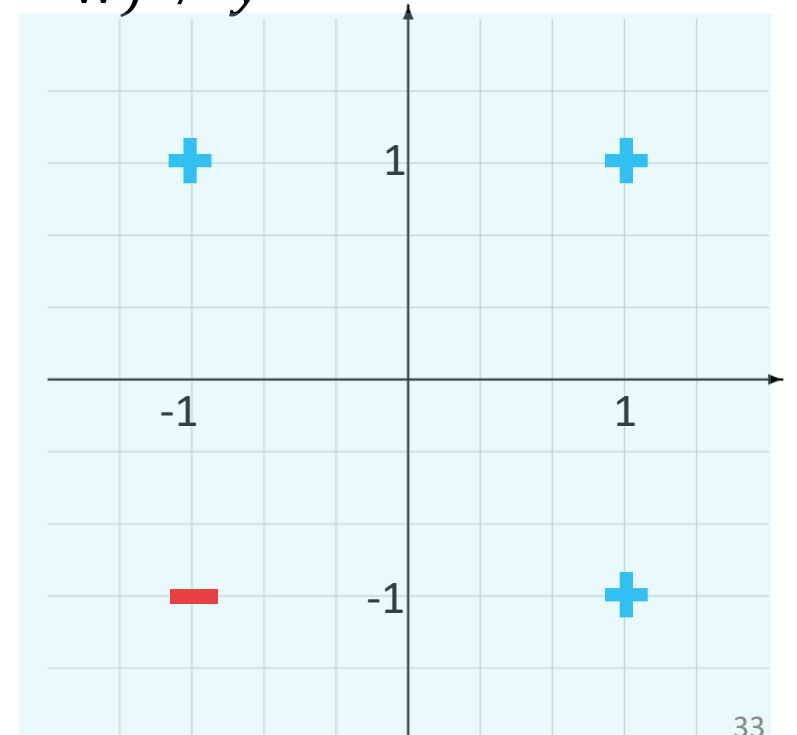
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - ▶ Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [2,0,0]$



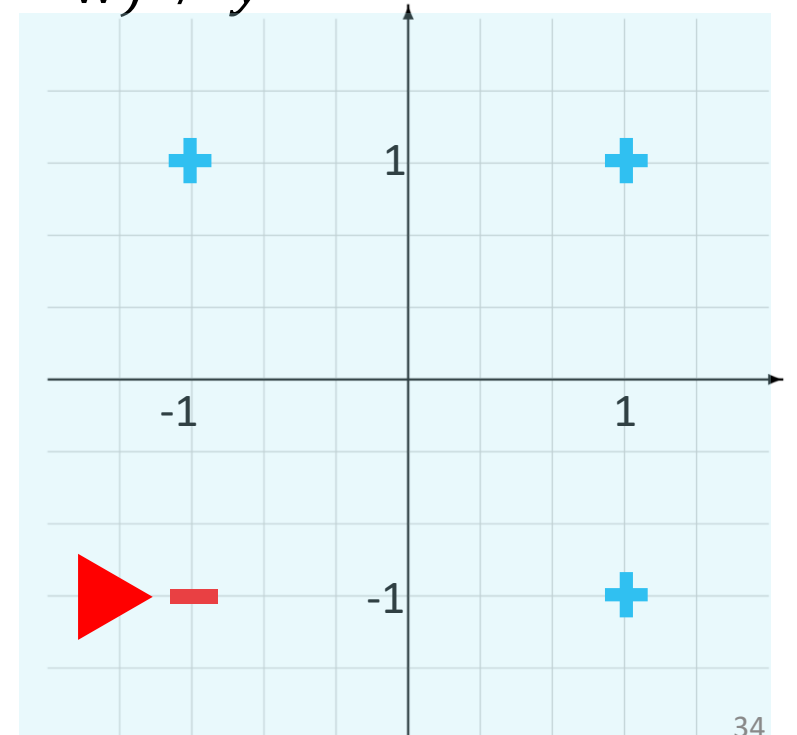
Exemplo

- Inicialize os pesos w aleatoriamente;
- ▶ Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [2,0,0]$



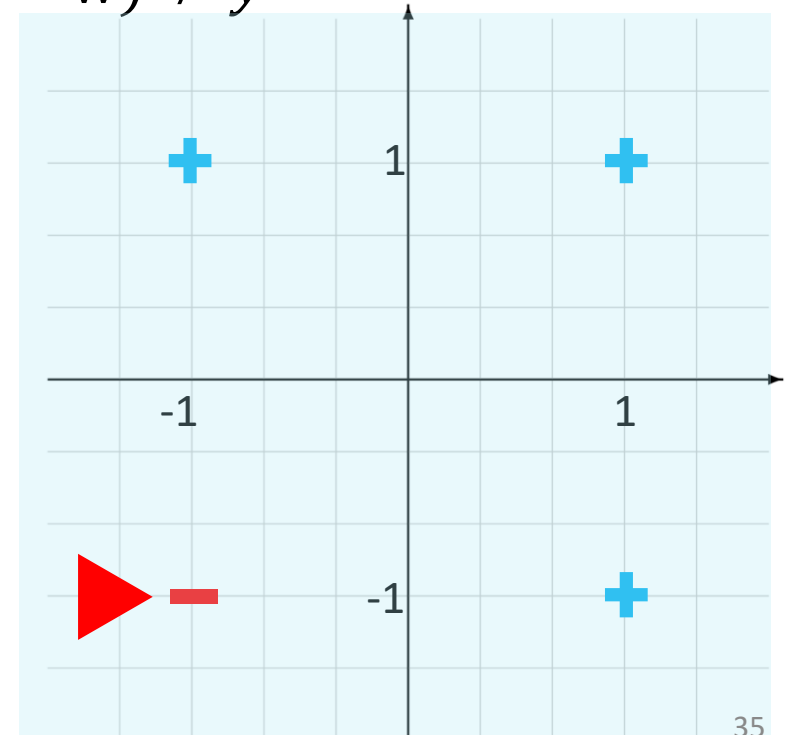
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - ▶ Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [2,0,0]$



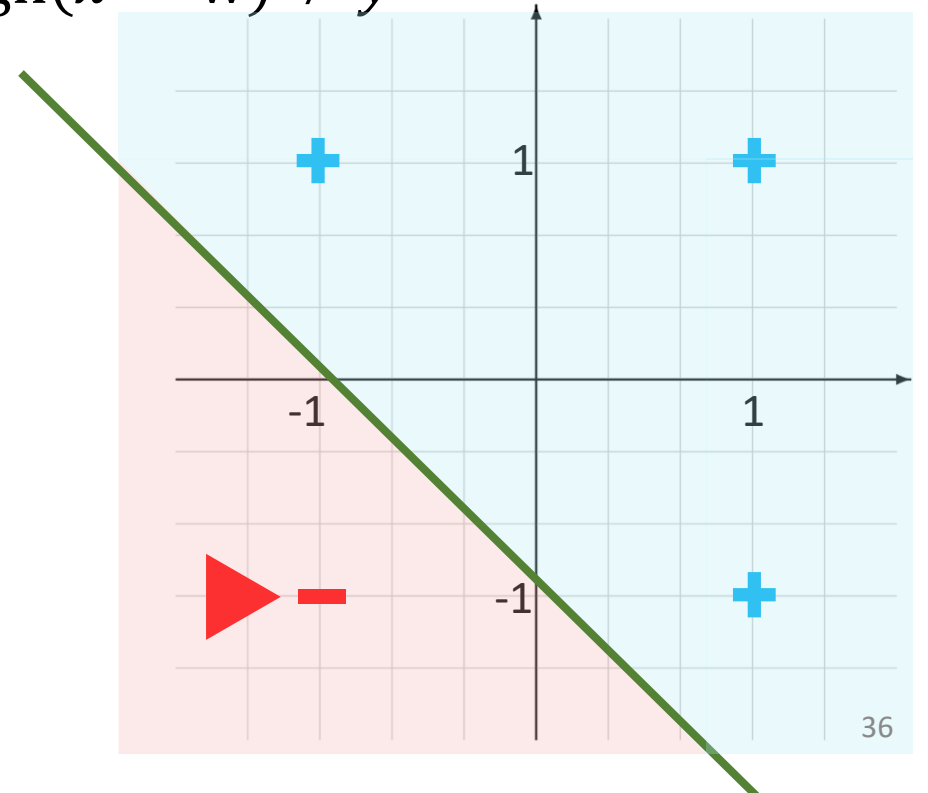
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - ▶ Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [2, 0, 0]$
- $w = [2, 0, 0] + (-1)[1, -1, -1]$



Exemplo

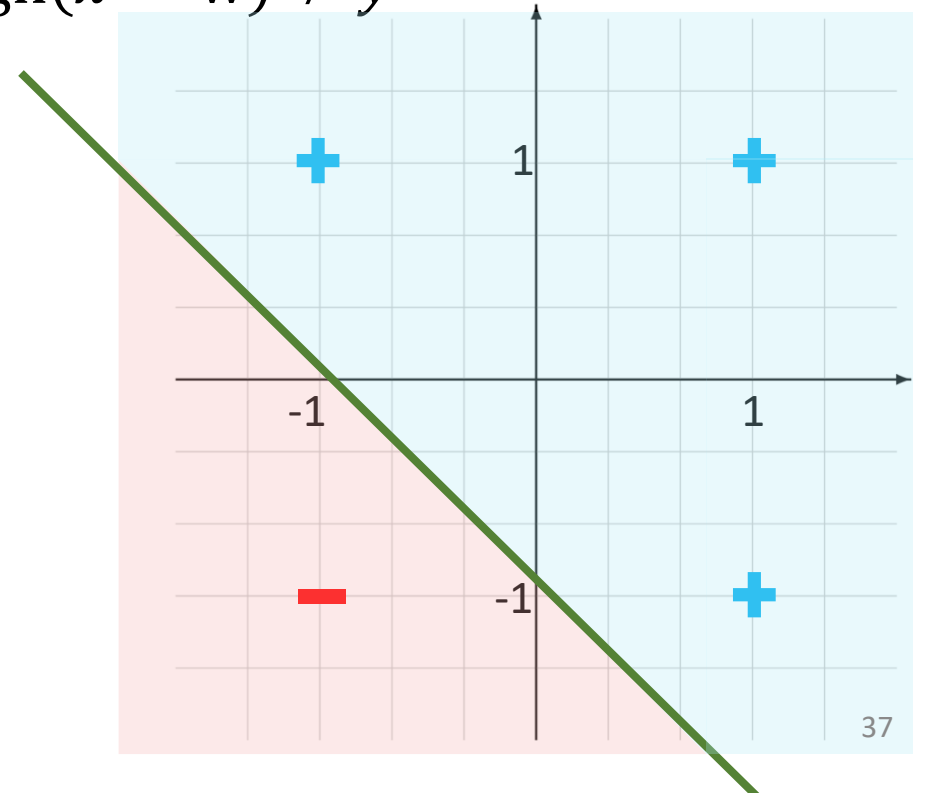
- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - ▶ Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [1,1,1]$



Exemplo

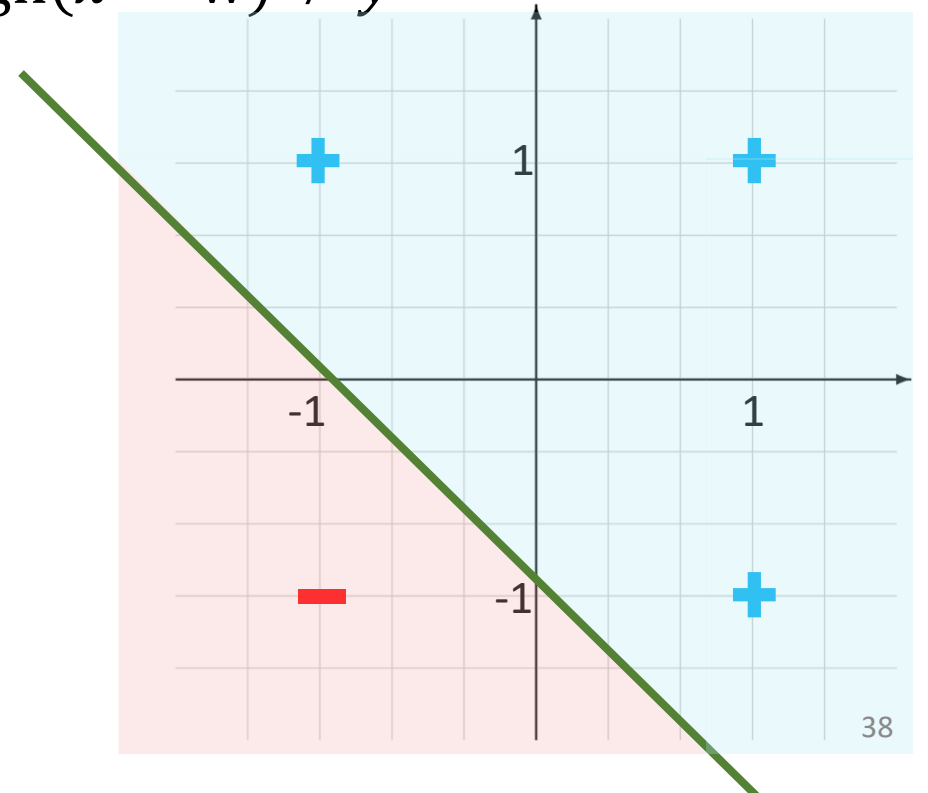
- Inicialize os pesos w aleatoriamente;
- ▶ Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$

- $w^T = [1,1,1]$



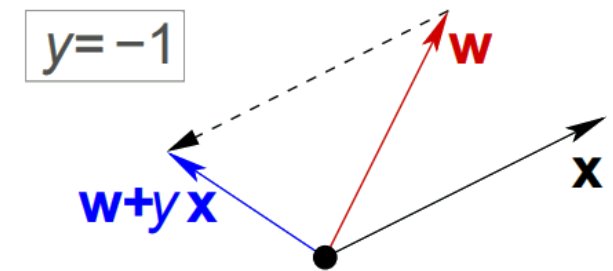
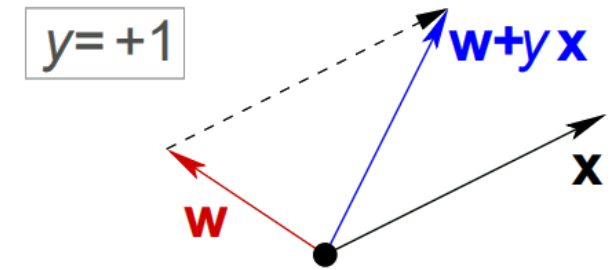
Exemplo

- Inicialize os pesos w aleatoriamente;
- Enquanto existirem observações classificadas erroneamente
 - Selecione uma observação $x^{(m)}$ tal que: $\text{sign}(x^{(m)}w) \neq y^{(m)}$
 - Atualize os pesos: $w = w + y^{(m)}x^{(m)}$
- $w^T = [1,1,1]$



Considerações sobre PLA

- É um algoritmo descoberto/criado de forma empírica
- De maneira intuitiva
 - A projeção de θ sobre as instâncias positivas devem ser positivas
 - A projeção de θ sobre as instâncias negativas devem ser negativas
- O PLA tenta atender a esses requisitos

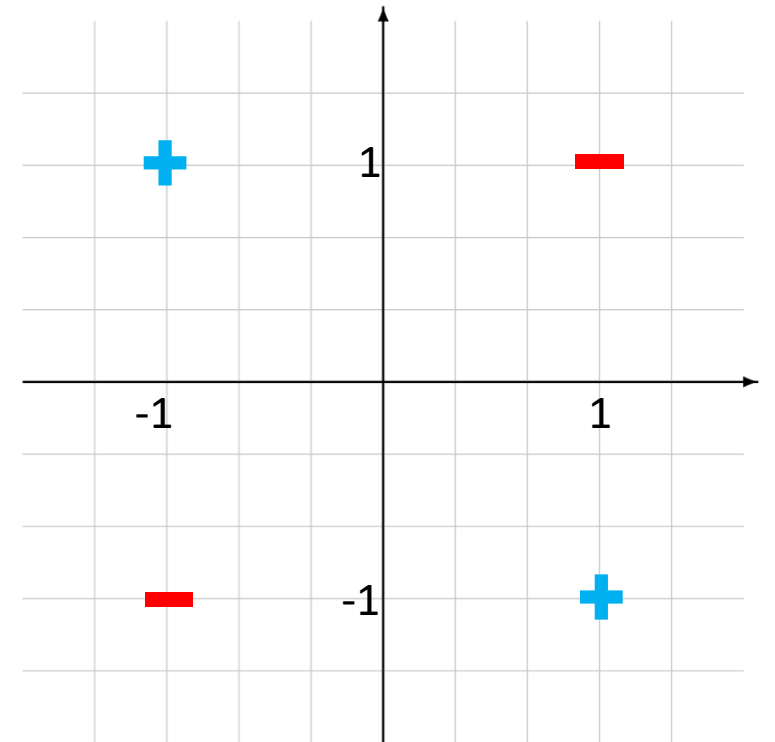


Considerações sobre PLA

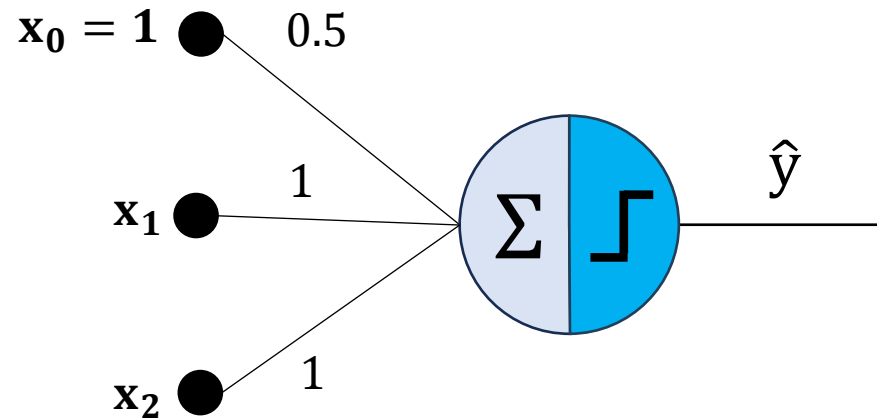
- Para casos não linearmente separáveis pode ser utilizado o “*Pocket Algorithm*”
- Armazenamos os melhores pesos durante o treinamento conforme a acurácia do perceptron

Limitação do Perceptron

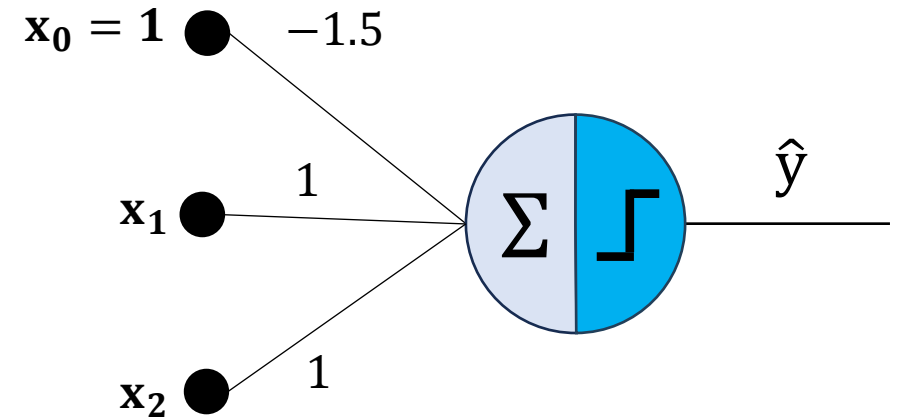
- Problemas não linearmente separáveis;
- XOR



Preencham as tabelas

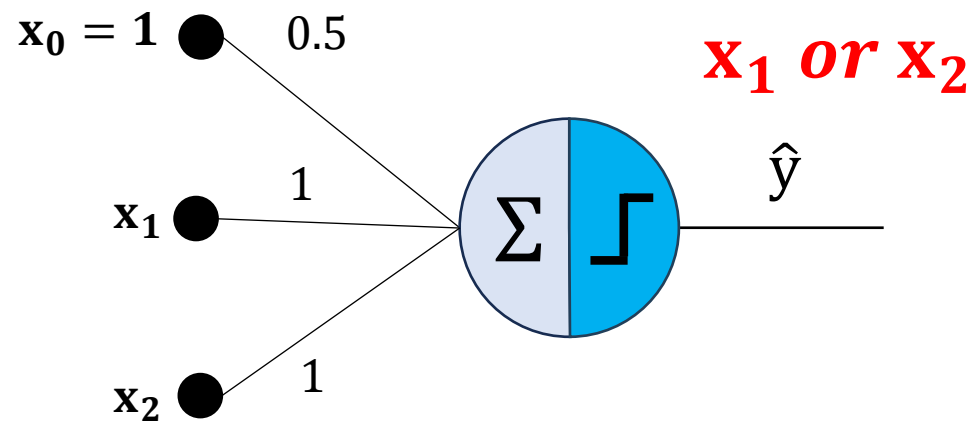


x_1	x_2	y
-1	-1	
-1	1	
1	-1	
1	1	

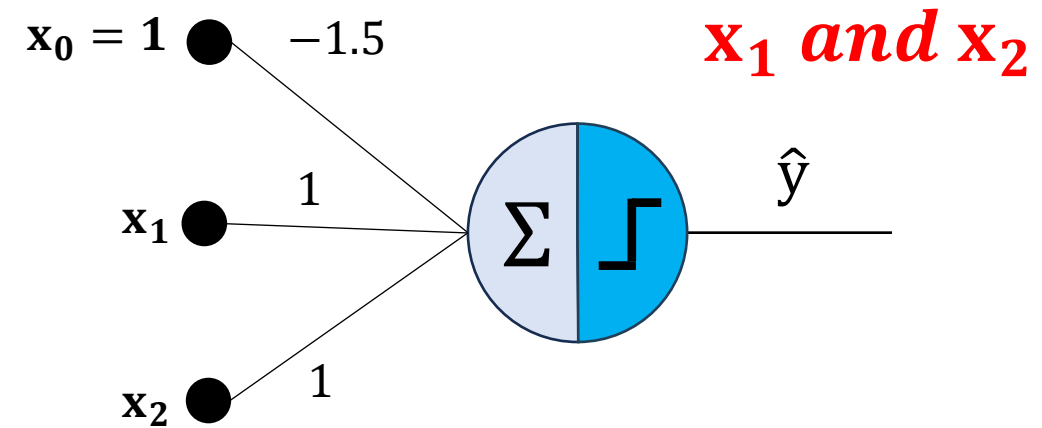


x_1	x_2	y
-1	-1	
-1	1	
1	-1	
1	1	

Operações lógicas implementadas

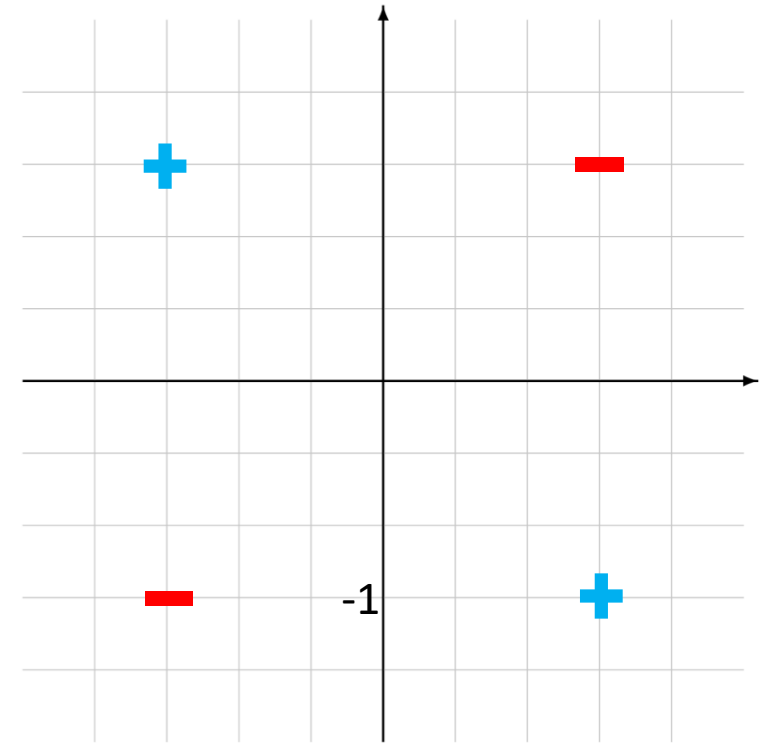
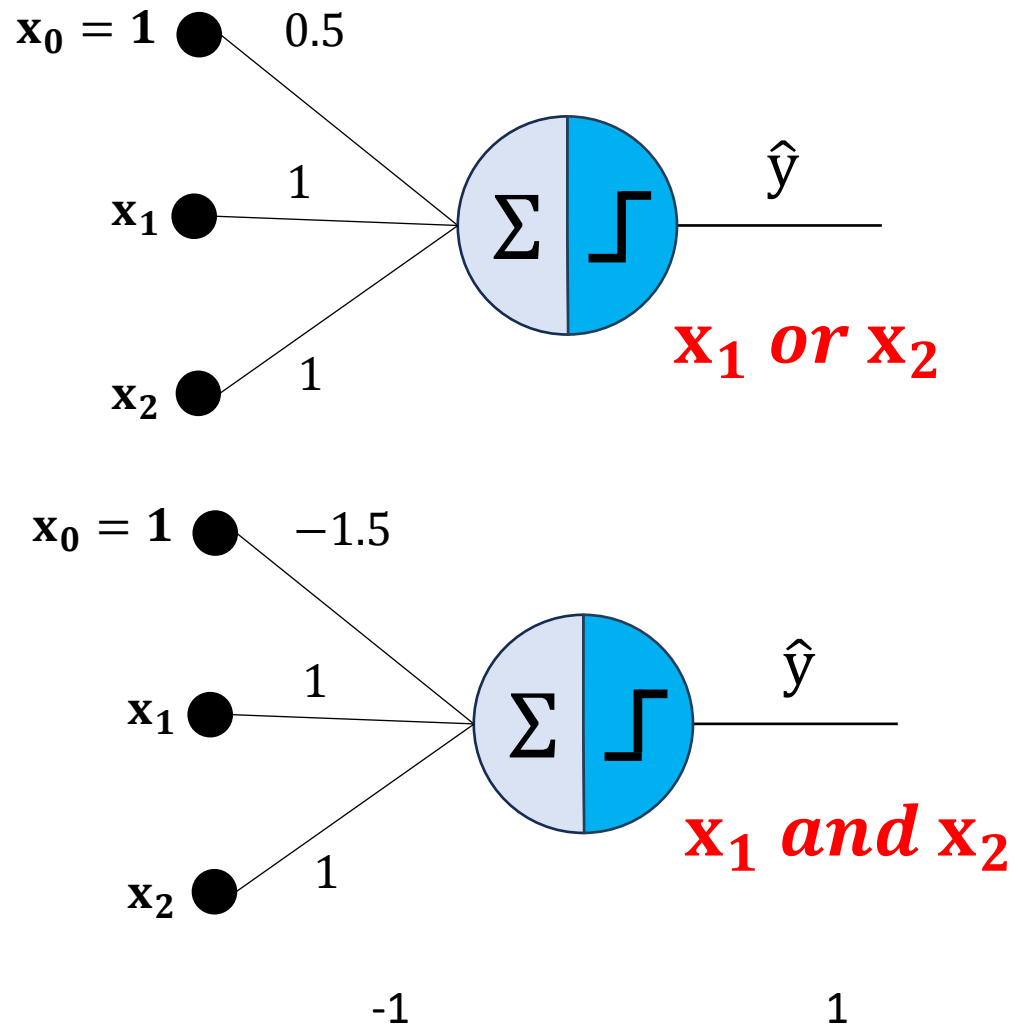


x_1	x_2	y
-1	-1	-1
-1	1	1
1	-1	1
1	1	1

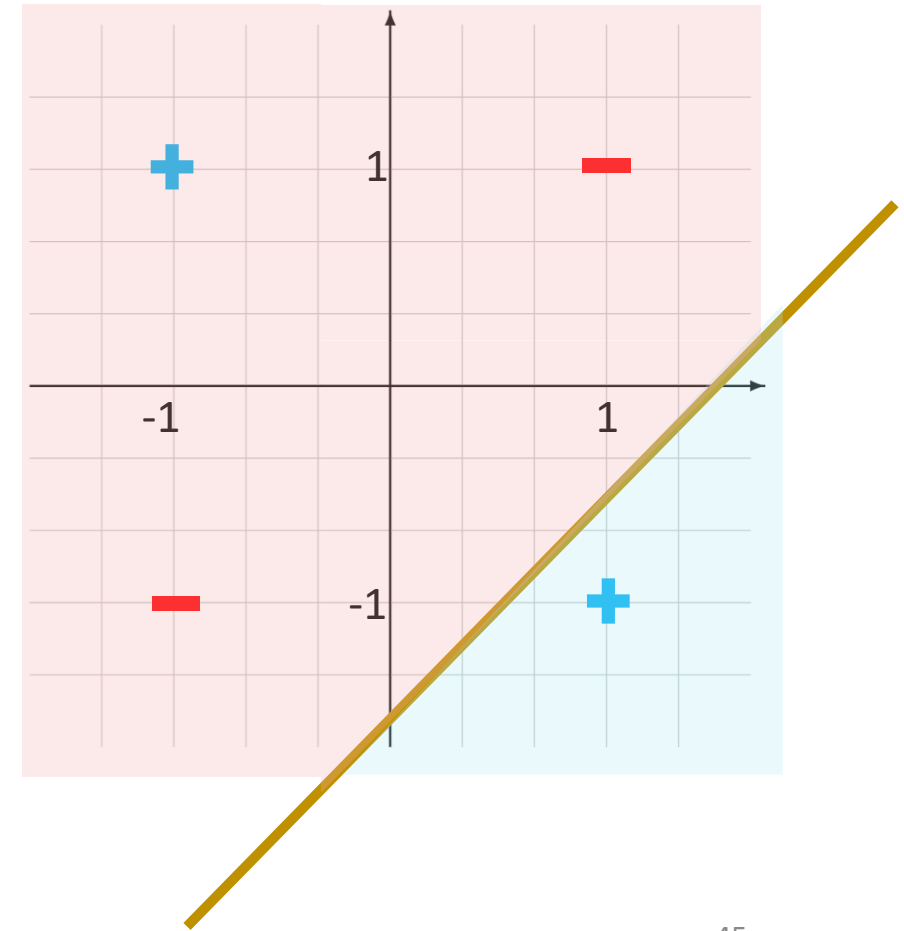
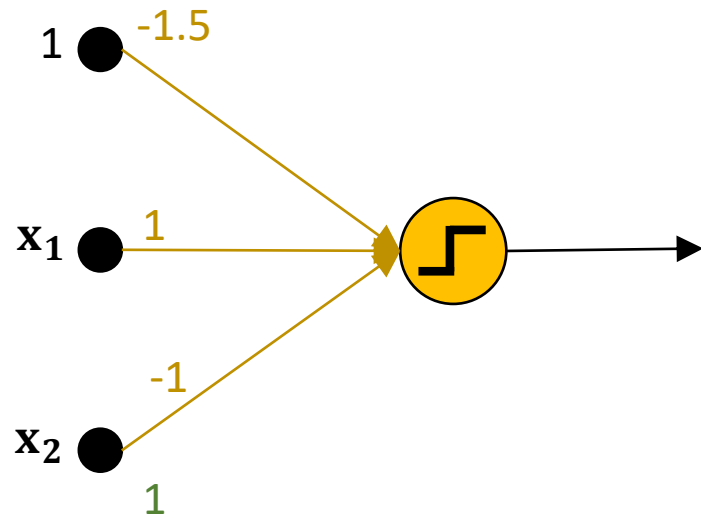


x_1	x_2	y
-1	-1	-1
-1	1	-1
1	-1	-1
1	1	1

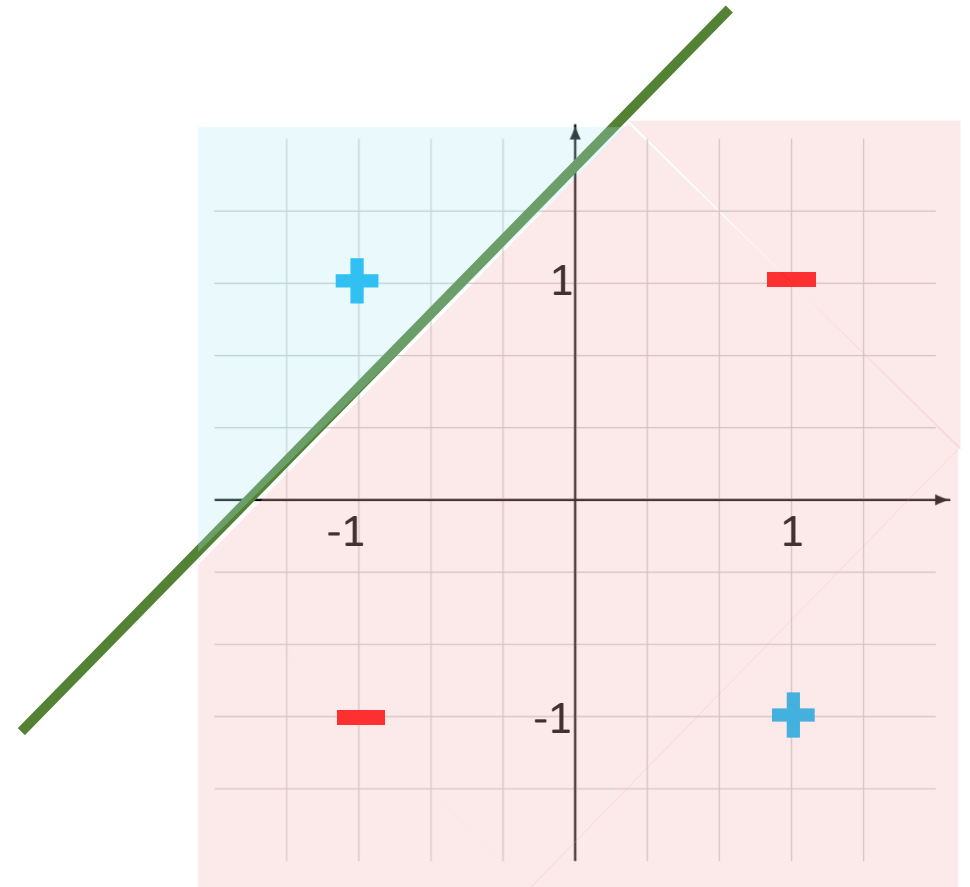
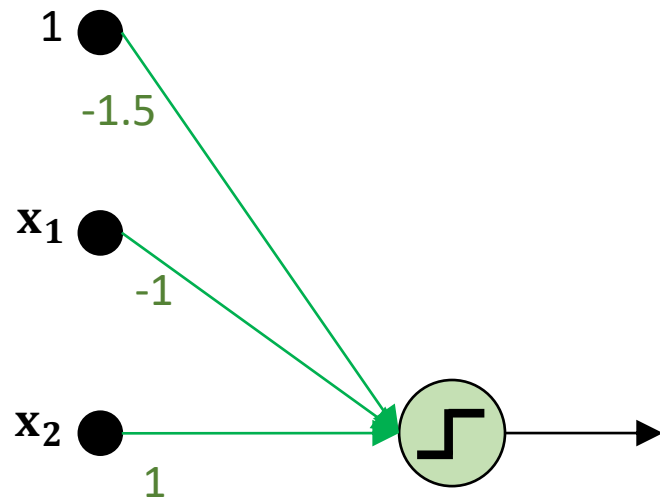
Podemos fazer algo a respeito?



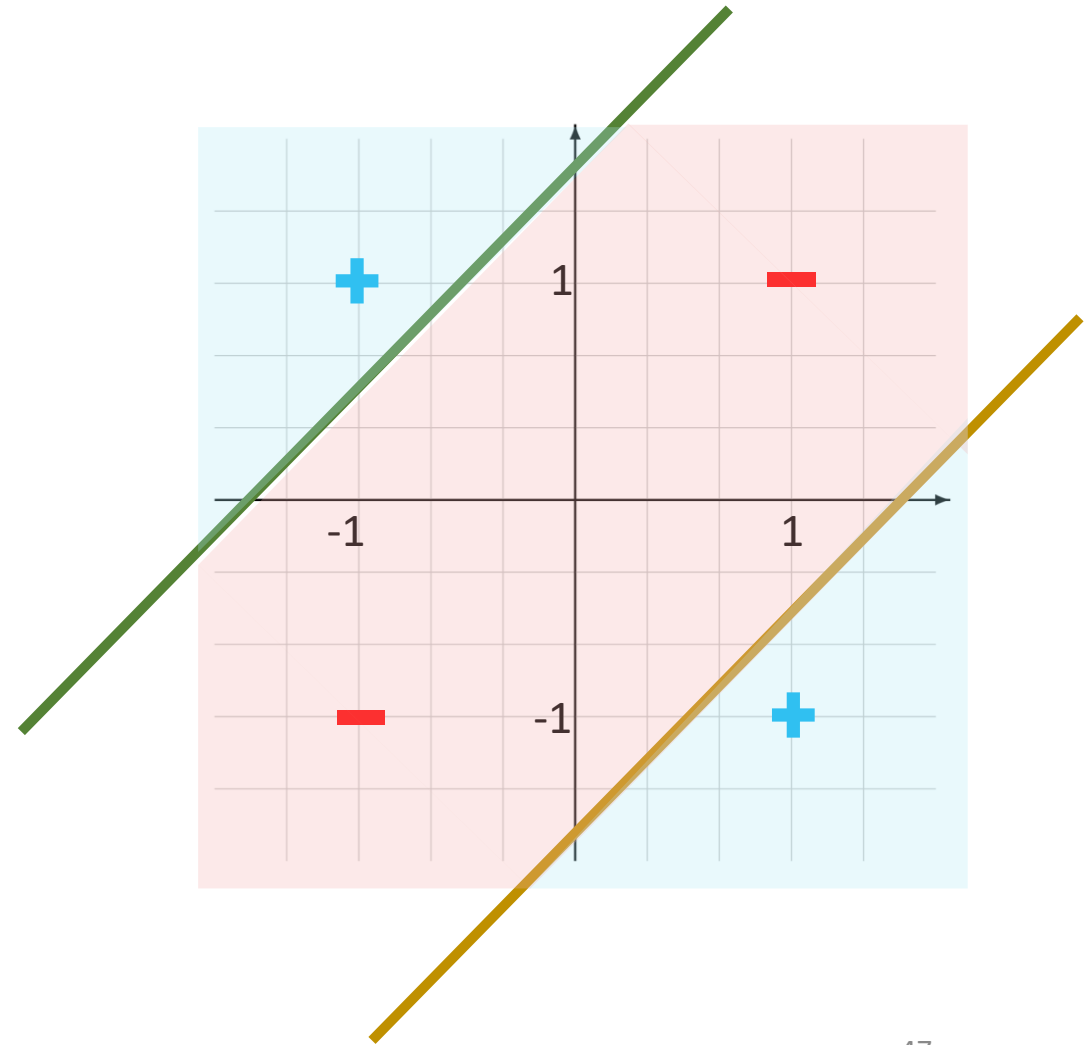
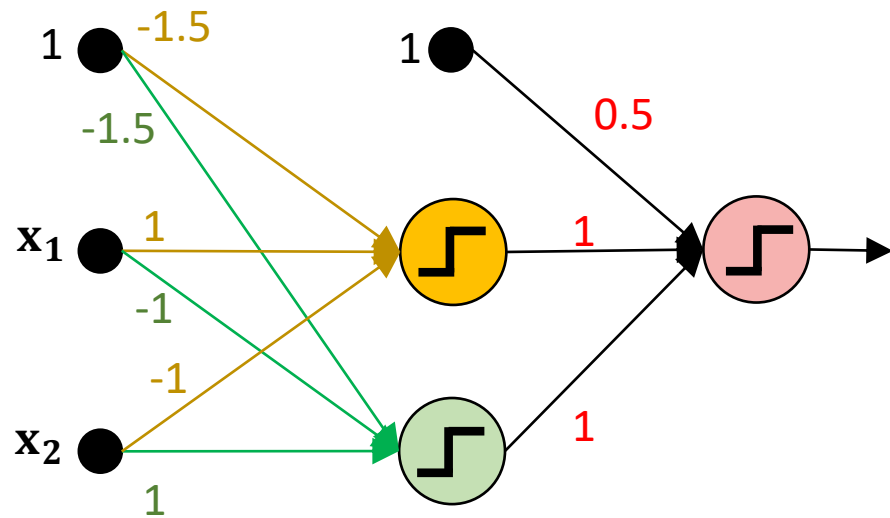
Podemos!



Podemos!



Podemos!



Leituras indicadas:

Para se aprofundar no conteúdo dessa aula:

- Capítulo 1 - ABU-MOSTAFA, Yaser S.; MAGDON-ISMAIL, Malik; LIN, Hsuan-Tien. **Learning from data**. New York, NY, USA:: AMLBook, 2012.
- Capítulo 1 - HAYKIN, Simon. **Neural networks and learning machines**, 3/E. Pearson Education India, 2010.

Preparação para as próximas aulas:

- Parte II – Capítulo 6 - GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep learning**. MIT press, 2016.
 - <https://www.deeplearningbook.org/contents/mlp.html>

Referências:

- HAYKIN, Simon. **Neural networks and learning machines, 3/E**. Pearson Education India, 2010.
- ABU-MOSTAFA, Yaser S.; MAGDON-ISMAIL, Malik; LIN, Hsuan-Tien. **Learning from data**. New York, NY, USA:: AMLBook, 2012.