



PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
ESCOLA POLITÉCNICA

MALTA

Machine Learning Theory
and Applications Lab

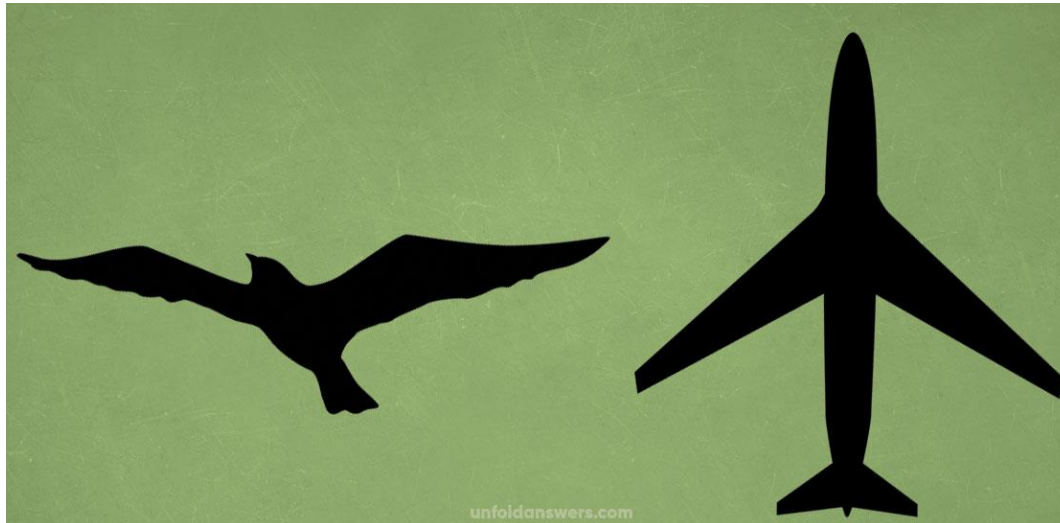
Aprendizado de Máquina

Paradigma baseado em Otimização
Redes Neurais I

Prof. Me. Otávio Parraga

Origem

- Para muitas invenções, nos baseamos na biologia



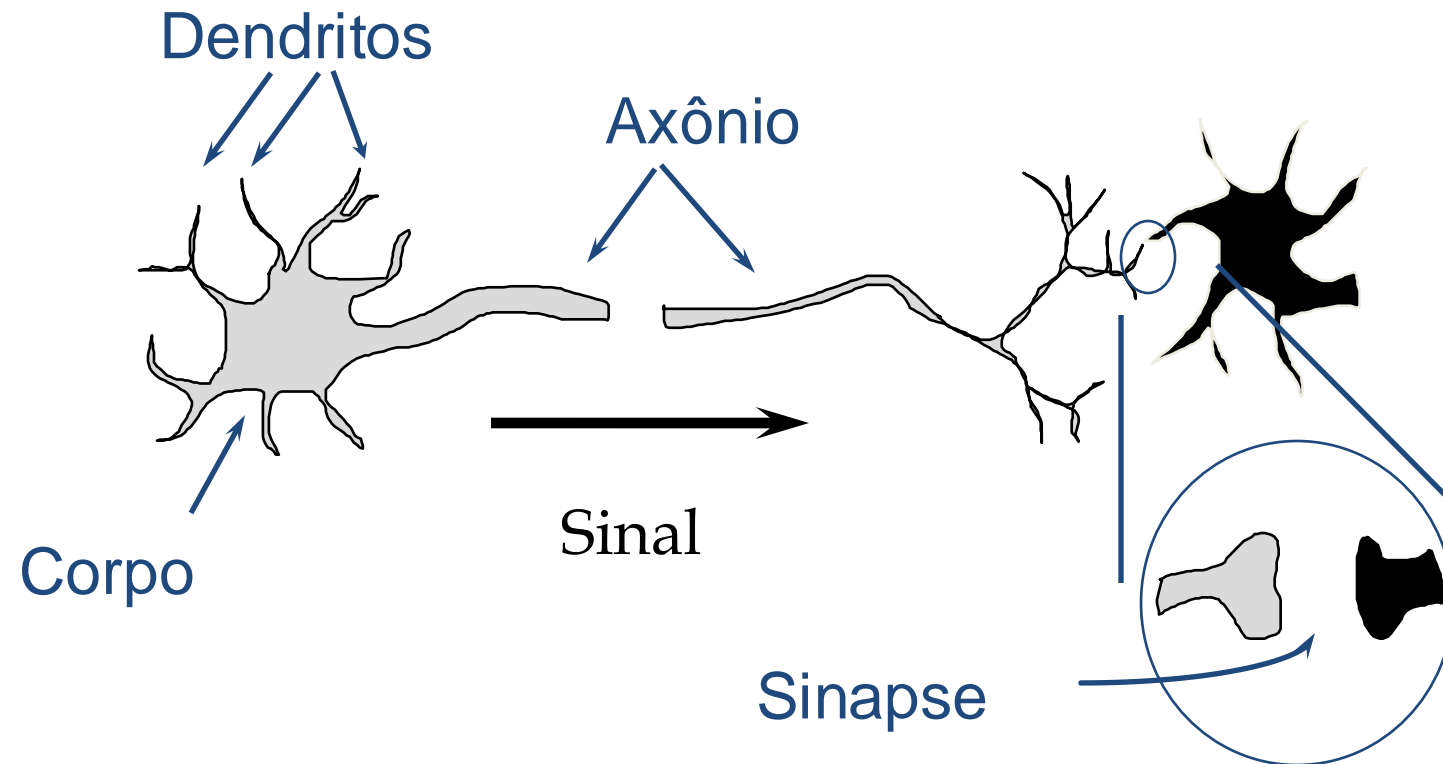
Origem

- Para muitas invenções, nos baseamos na biologia



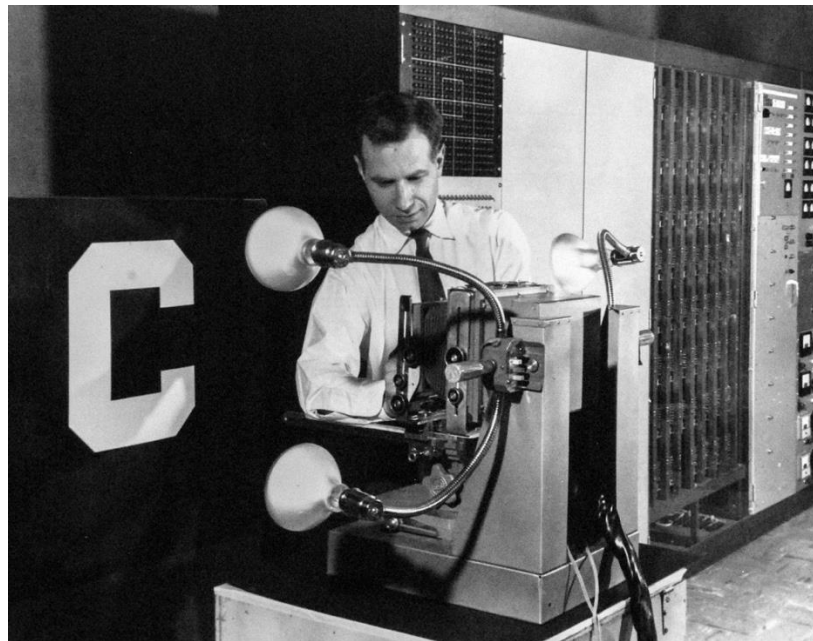
Origem

- Qual o melhor exemplo de inteligência na natureza?



Origem

- Formulado matematicamente por McCulloch e Pitts (1943)
- Desenvolvido por Rosenblatt (1958)

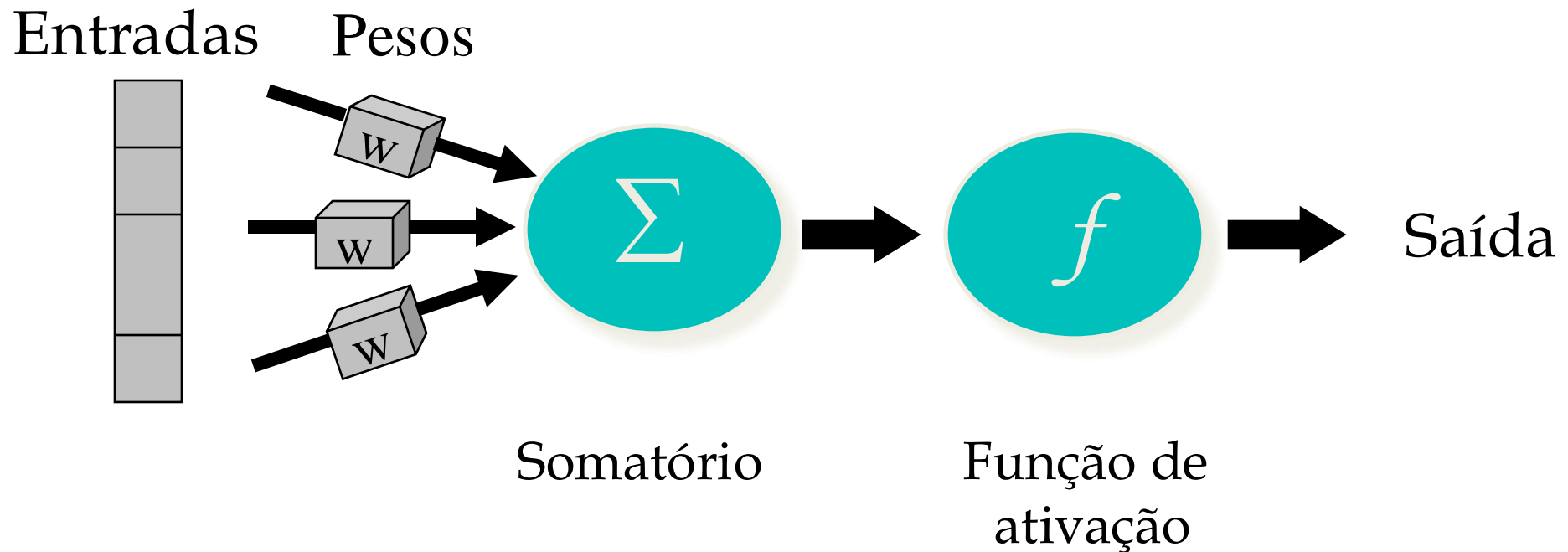


Origem



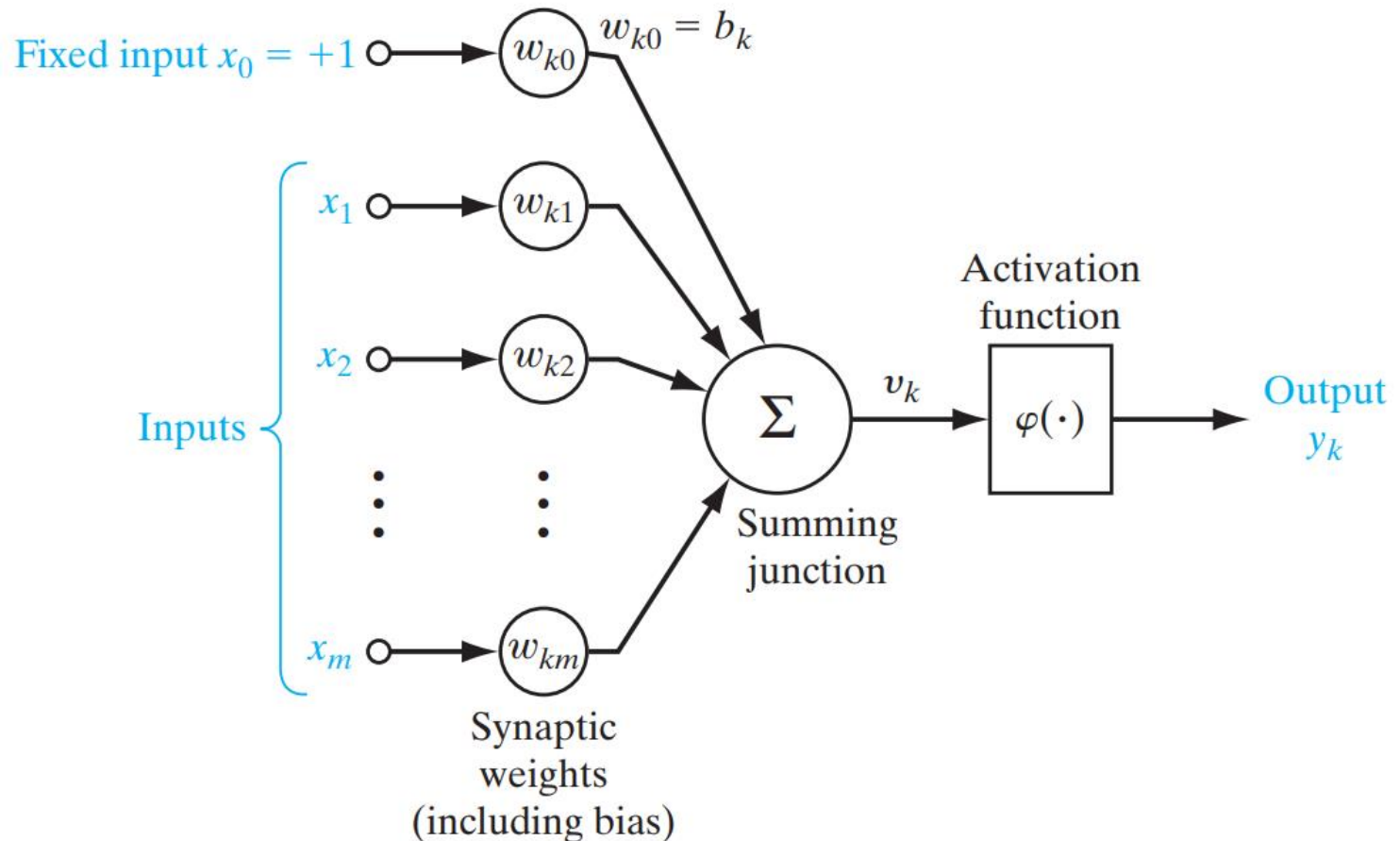
Estrutura do Neurônio

- Paradigma de Aprendizado Supervisionado



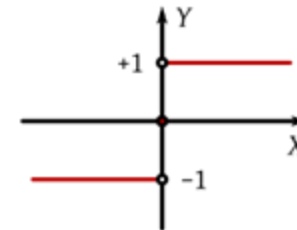
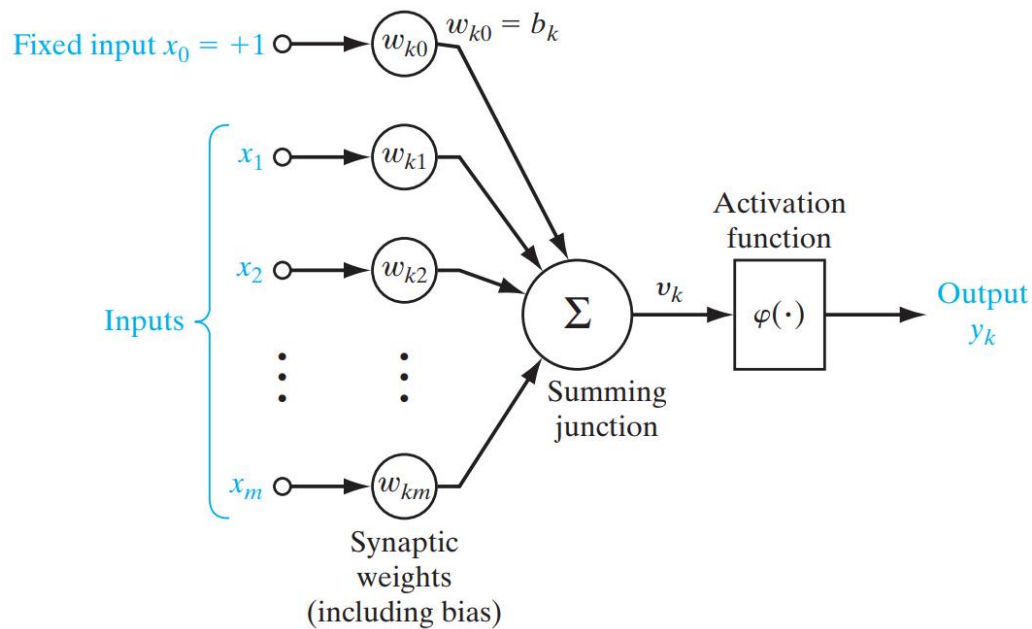
Estrutura do Neurônio

- Lembrar do termo independente (*bias*)



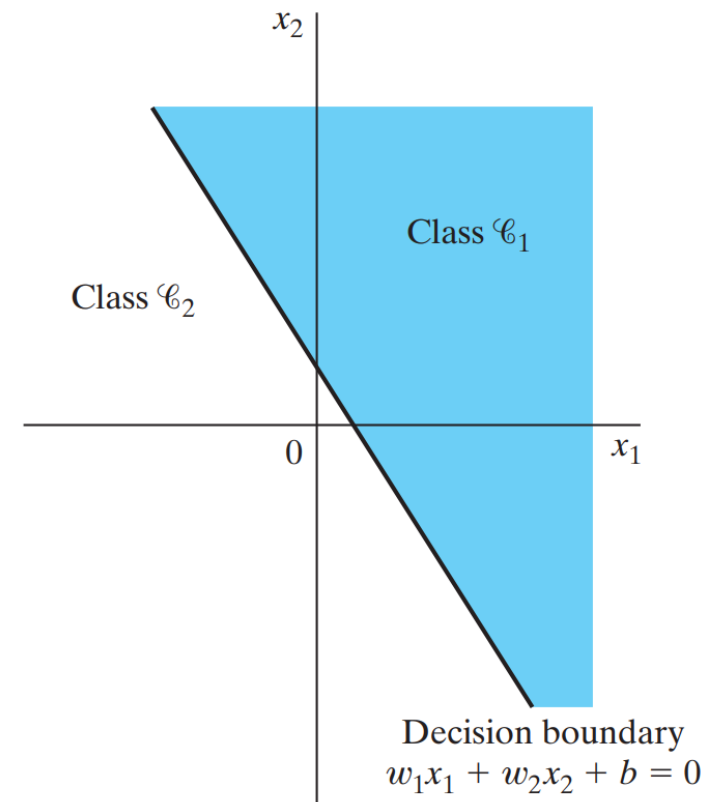
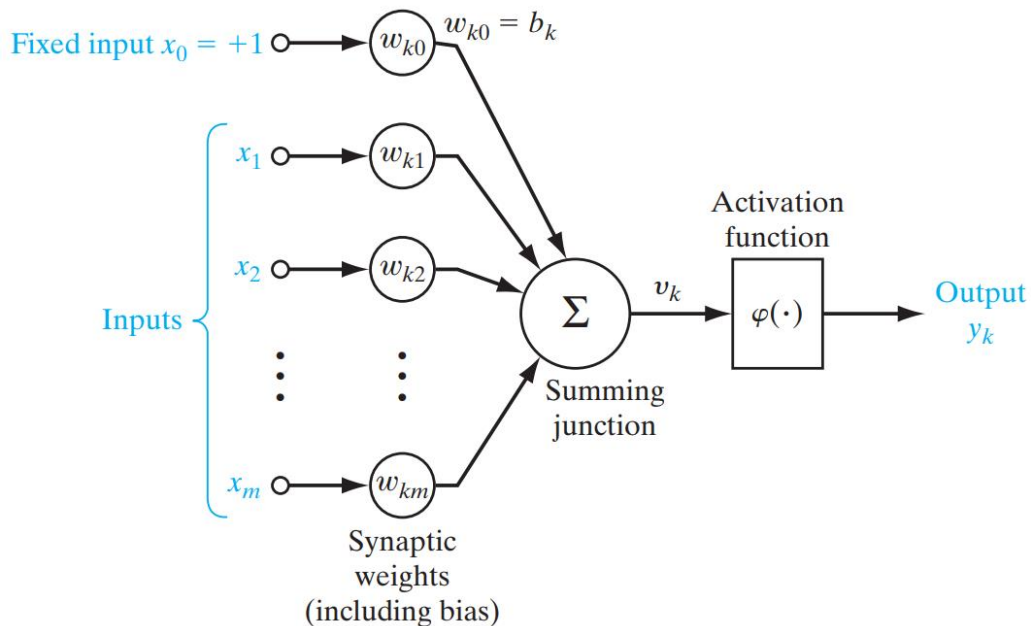
O Perceptron

- Função de Ativação -> Função Sinal
- $\hat{y} = \text{sign}(Xw)$
 - 1 ou -1



O Perceptron

- Fronteira de decisão linear



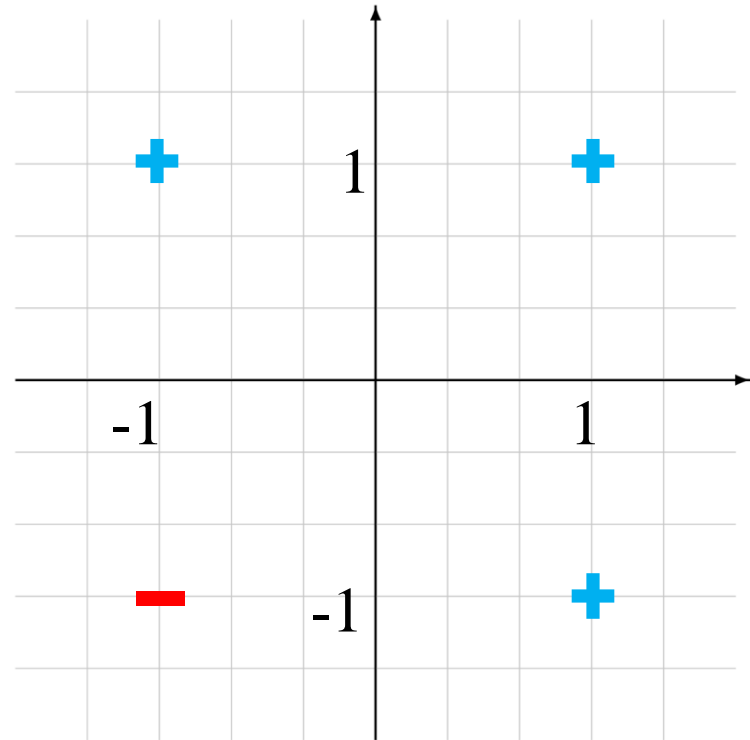
O Perceptron

- Como o Perceptron era treinado?
 - Perceptron Learning Algorithm (PLA)
 - Inicializar os pesos w aleatoriamente
 - Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

0 Perceptron

- Exemplo

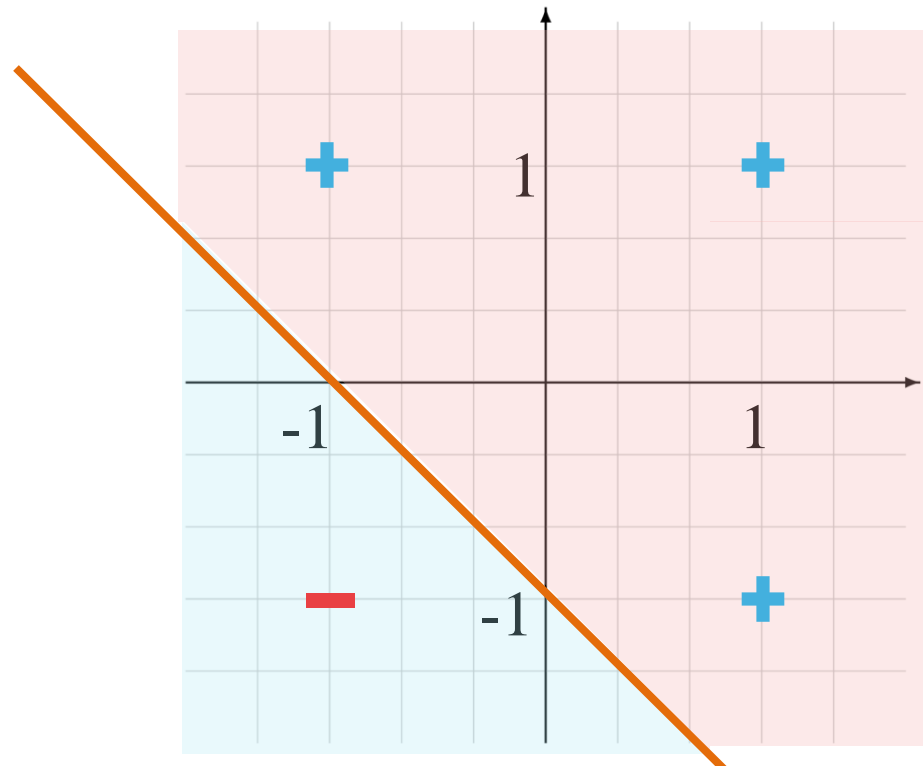
x_1	x_2	y
-1	-1	-1
-1	1	1
1	-1	1
1	1	1



Exemplo

- ▶ Inicializar os pesos w aleatoriamente
 - Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

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



Exemplo

- Como nós sabemos a fronteira de decisão:

- $EqReta = b + w_1x_1 + w_2x_2$

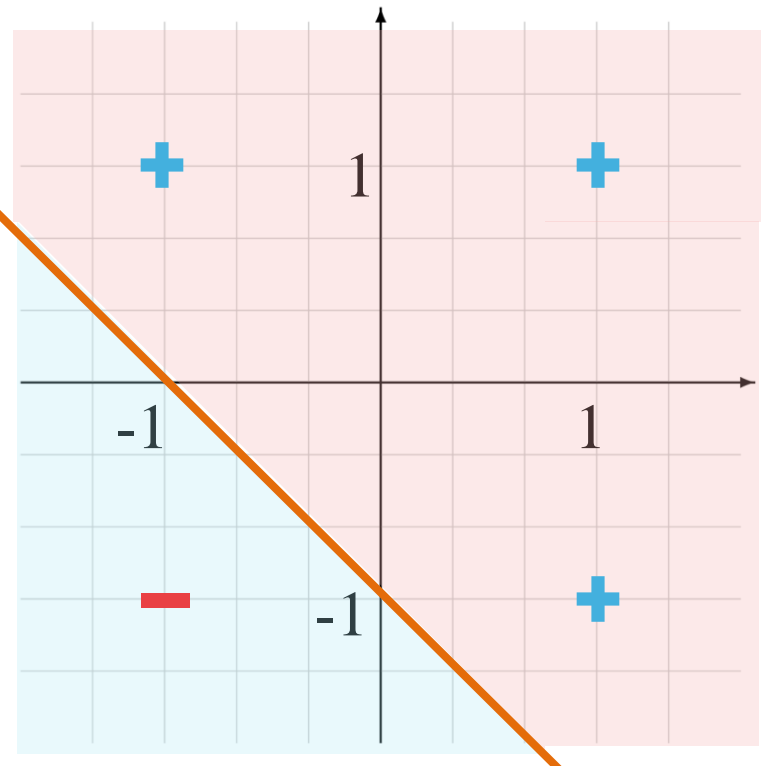
- $b + w_1x_1 + w_2x_2 = 0$

- $w_2x_2 = -w_1x_1 - b$

- $x_2 = -\frac{w_1x_1}{w_2} - \frac{b}{w_2}$

- $coef = -\frac{w_1x_1}{w_2}$

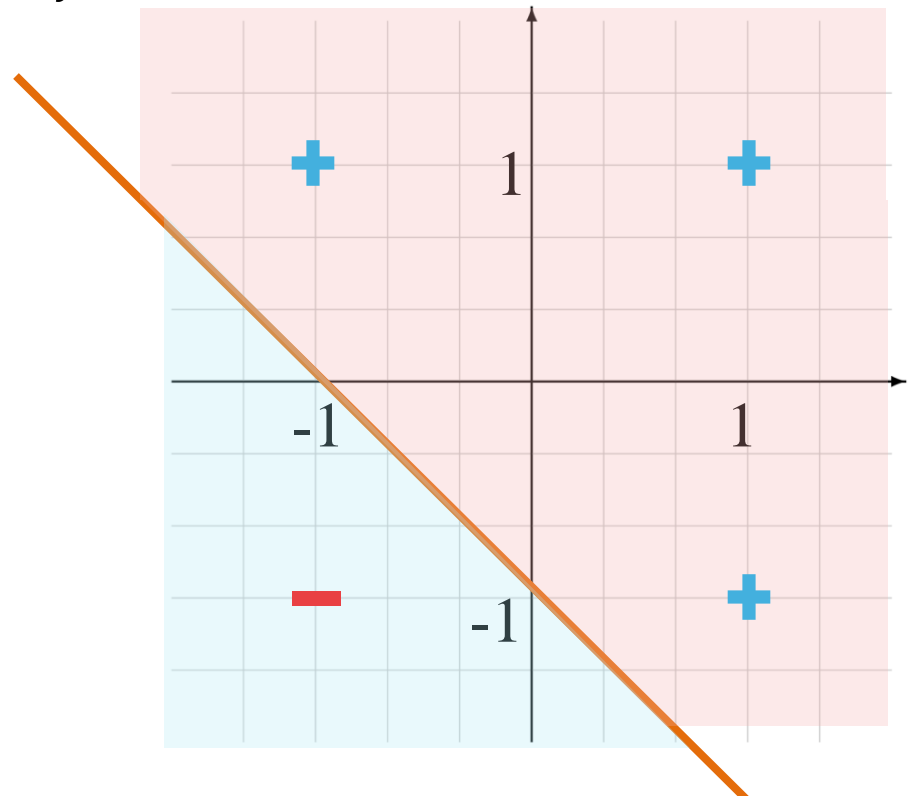
- $intercept = -\frac{b}{w_2}$



Exemplo

- ▶ Inicializar os pesos w aleatoriamente
 - Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

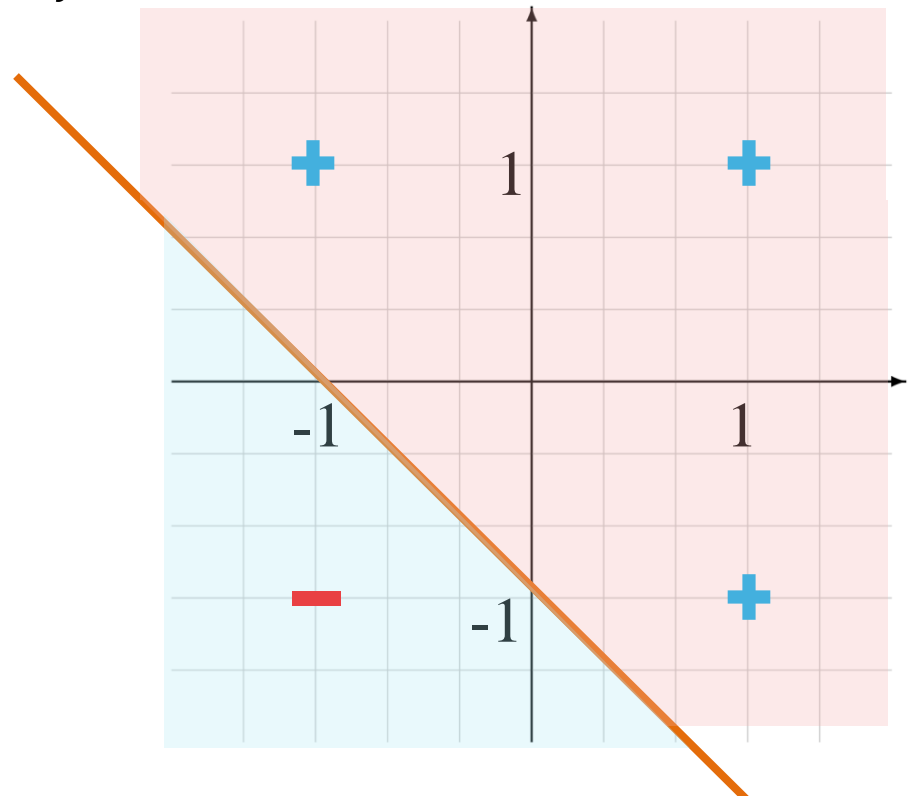
$$w^T = [-1, -1, -1]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- ▶ Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

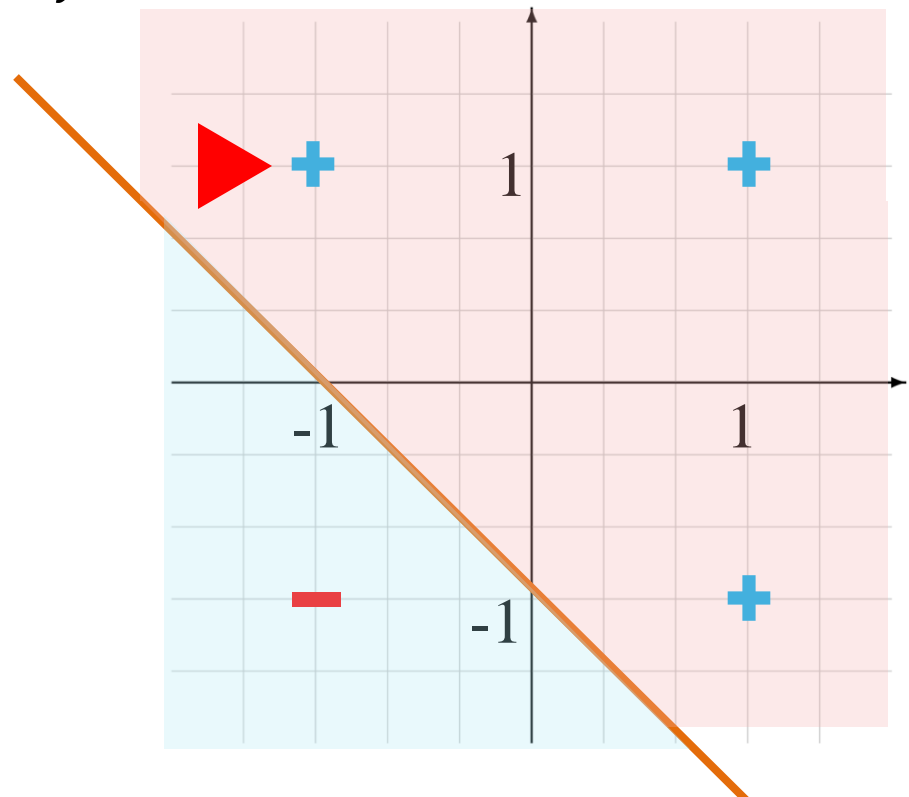
$$w^T = [-1, -1, -1]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- Enquanto existirem observações sendo classificadas de forma errada:
 - ▶ Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

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

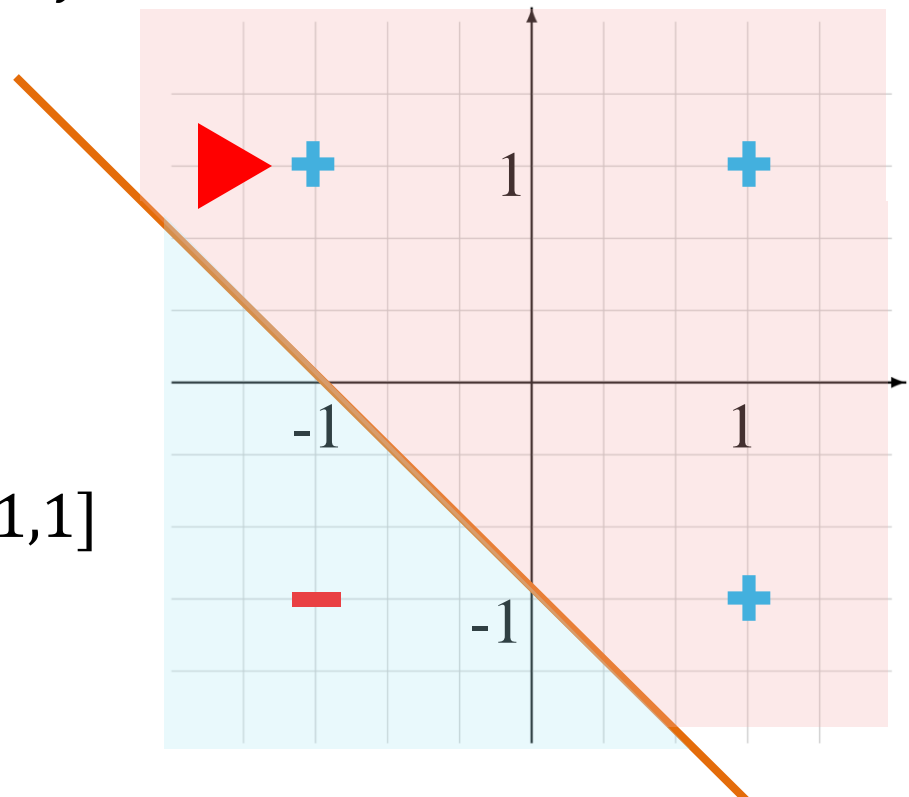


Exemplo

- Inicializar os pesos w aleatoriamente
- Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - ▶ Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

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

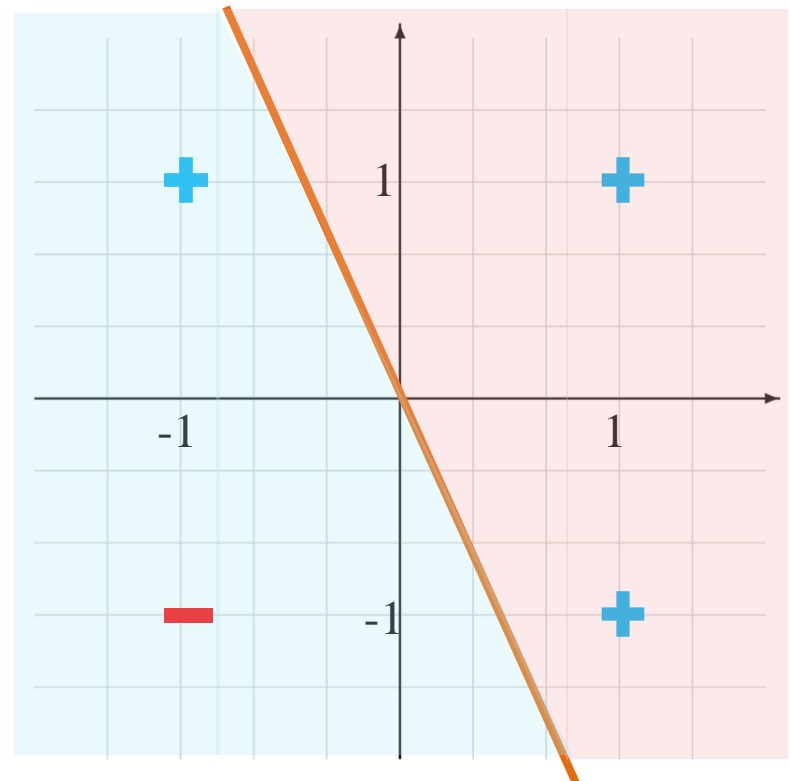
$$w = [-1, -1, -1] + [1, -1, 1]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- ▶ Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

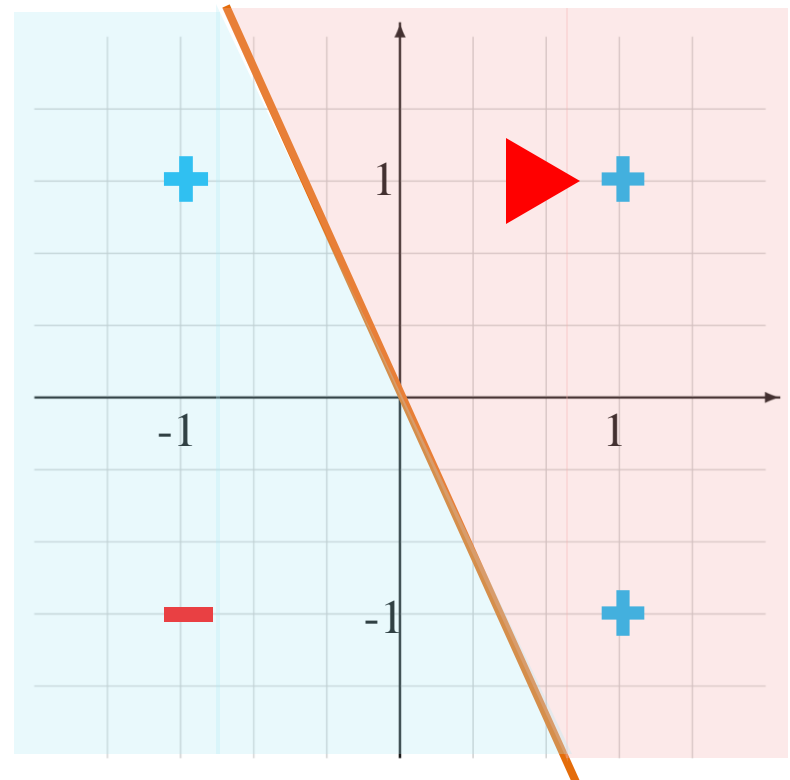
$$w^T = [0, -2, 0]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- Enquanto existirem observações sendo classificadas de forma errada:
 - ▶ Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

$$w^T = [0, -2, 0]$$

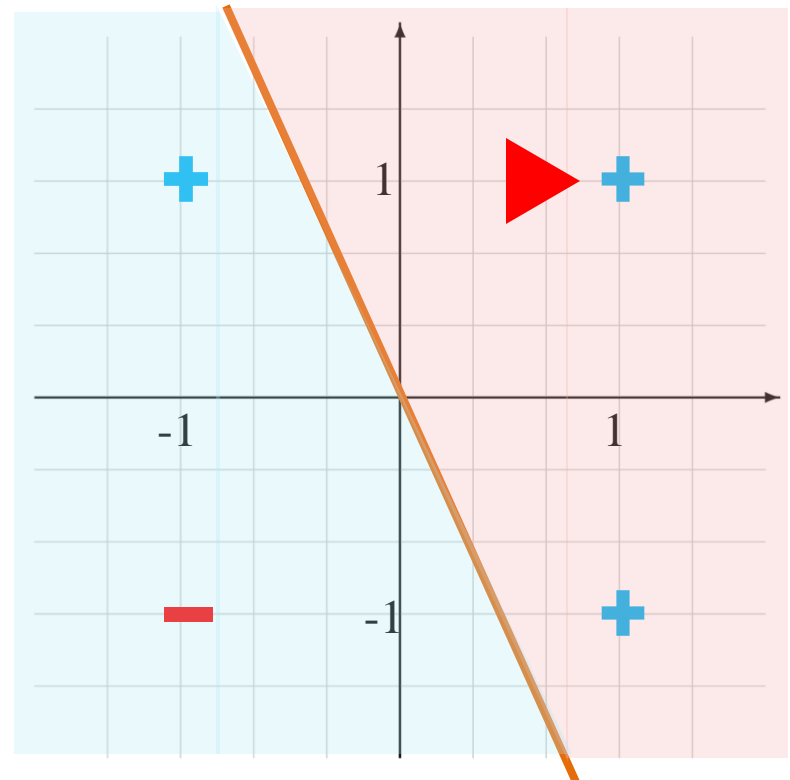


Exemplo

- Inicializar os pesos w aleatoriamente
- Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - ▶ Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

$$w^T = [0, -2, 0]$$

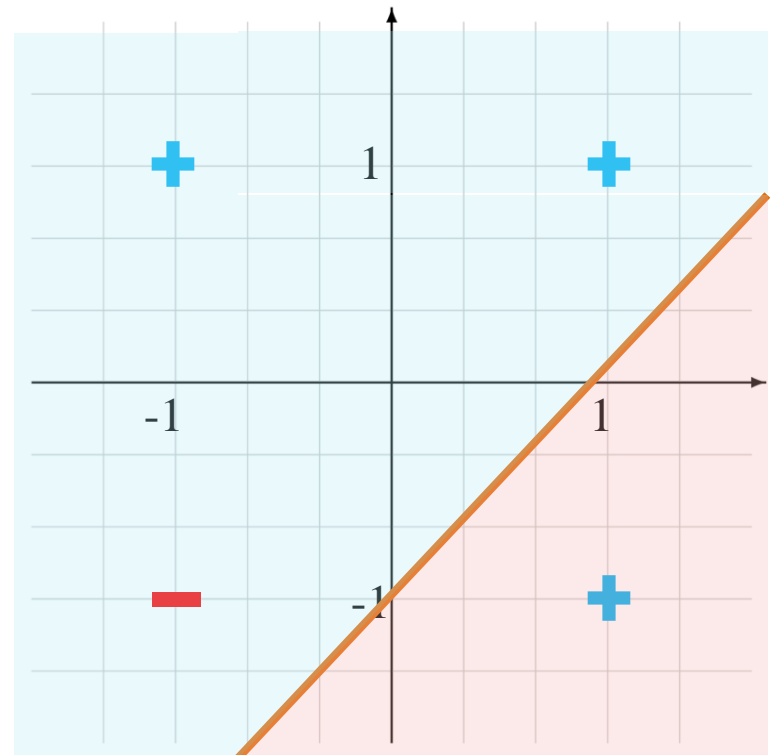
$$w = [0, -2, 0] + [1, 1, 1]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- ▶ Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

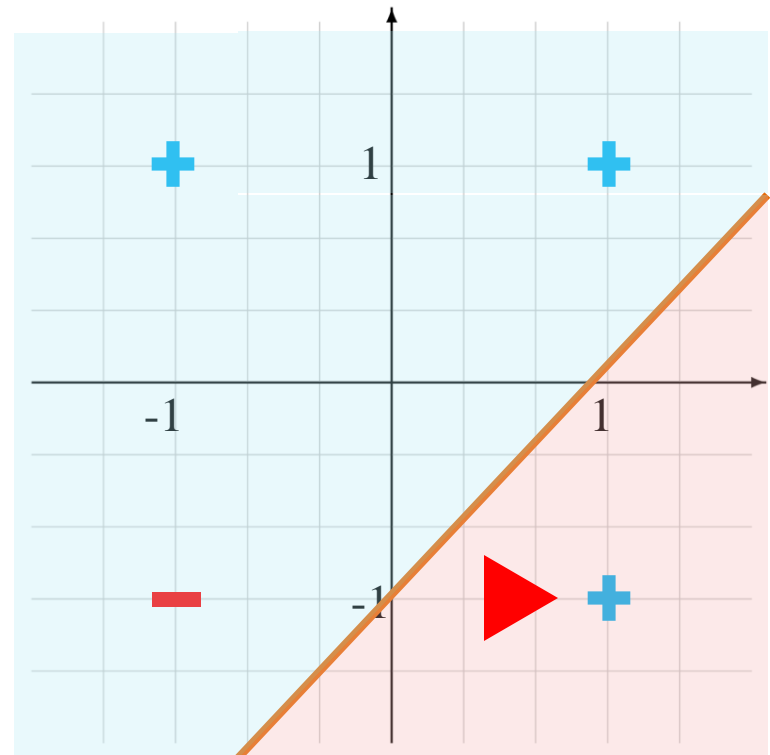
$$w^T = [1, -1, 1]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- Enquanto existirem observações sendo classificadas de forma errada:
 - ▶ Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

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

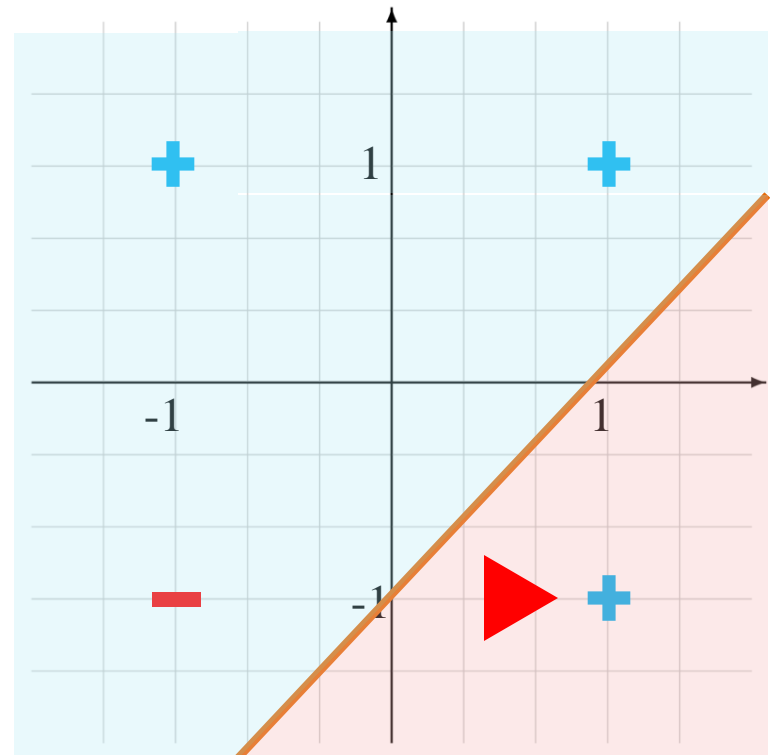


Exemplo

- Inicializar os pesos w aleatoriamente
 - Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
- Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

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

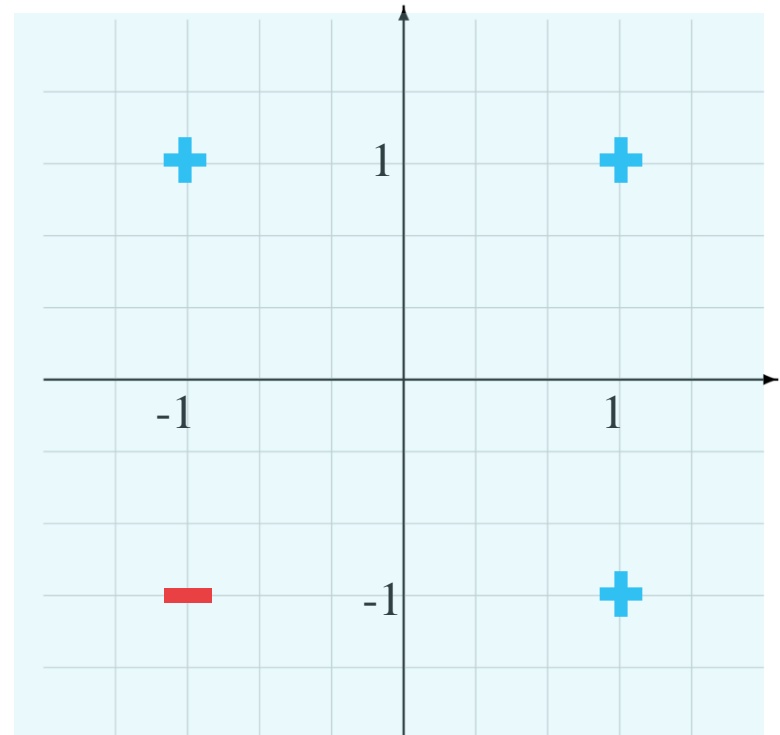
$$w = [1, -1, 1] + [1, 1, -1]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- ▶ Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

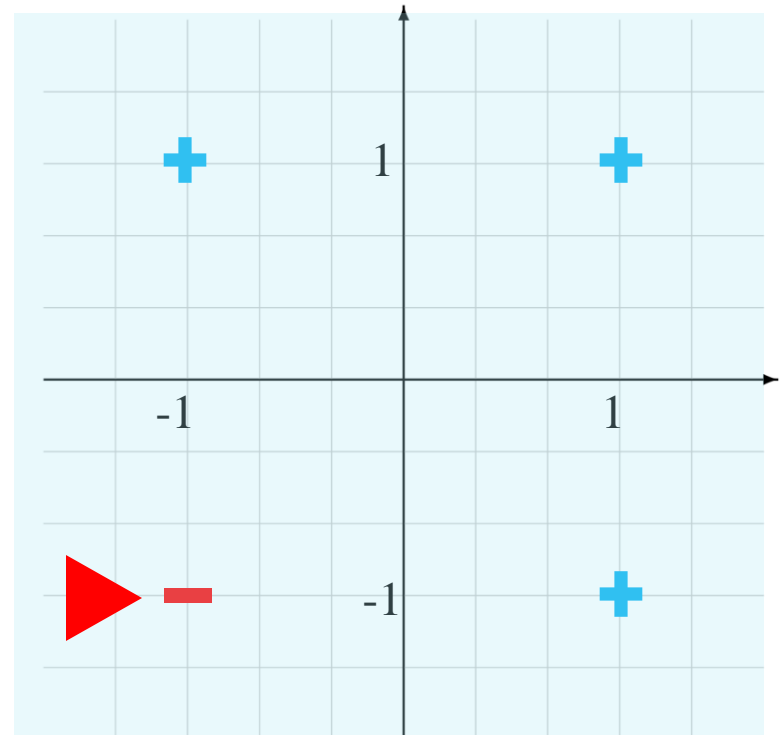
$$w^T = [2, 0, 0]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- Enquanto existirem observações sendo classificadas de forma errada:
 - ▶ Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}r^{(n)}$

$$w^T = [2, 0, 0]$$

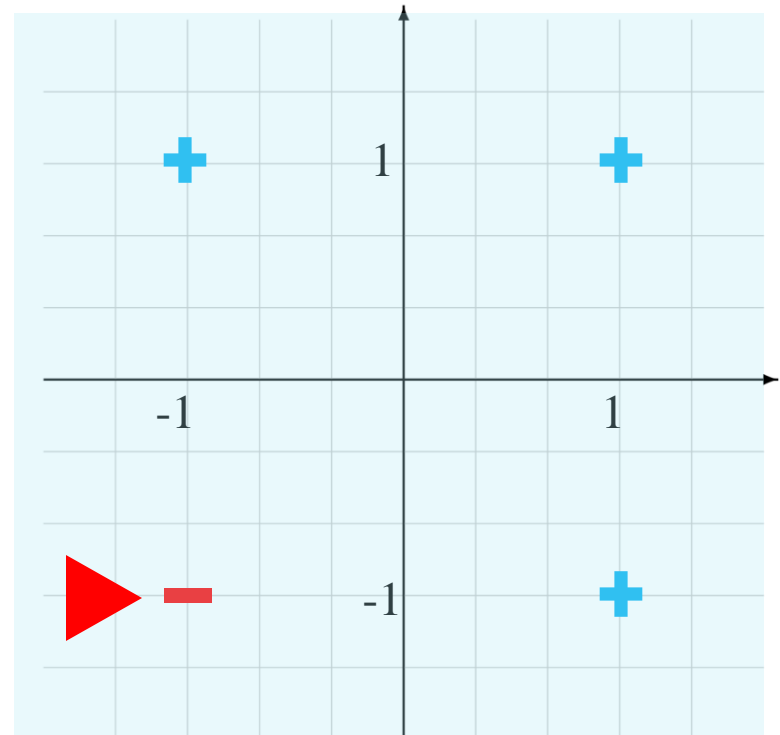


Exemplo

- Inicializar os pesos w aleatoriamente
 - Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
- Atualizar os pesos: $w = w + y^{(n)}r^{(n)}$

$$w^T = [2, 0, 0]$$

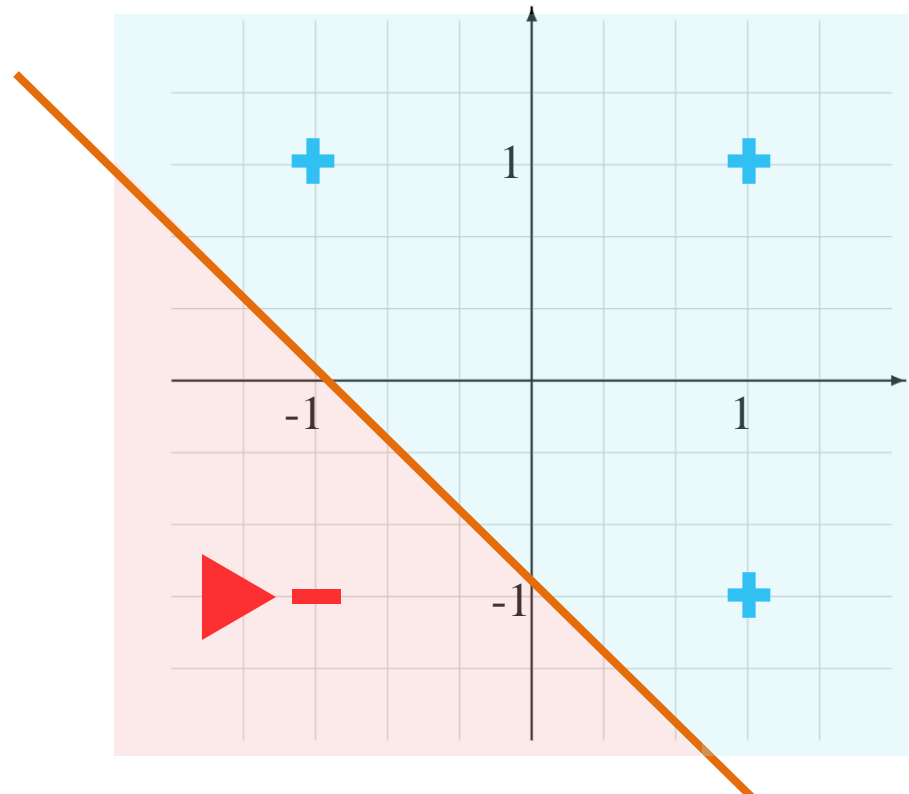
$$w = [2, 0, 0] + (-1)[1, -1, -1]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - ▶ Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

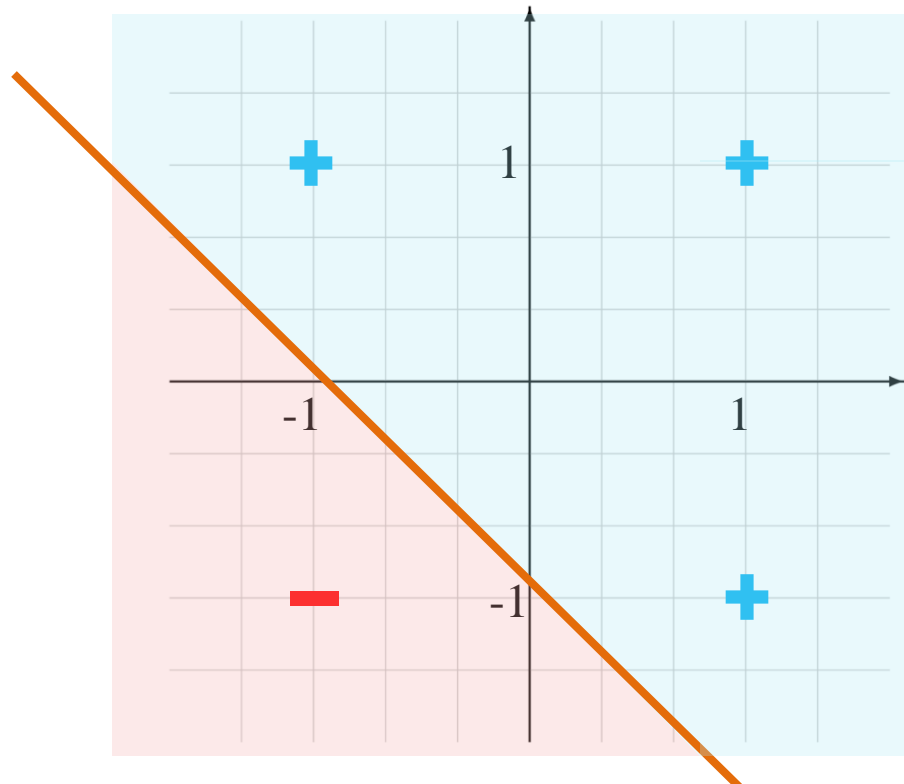
$$w^T = [1, 1, 1]$$



Exemplo

- Inicializar os pesos w aleatoriamente
- ▶ Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

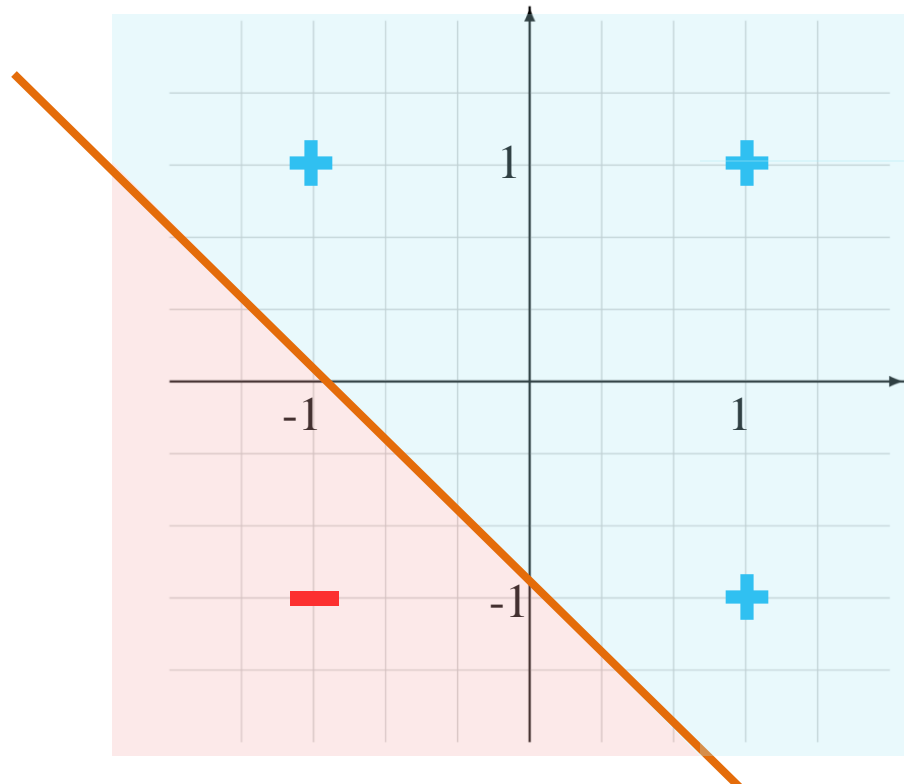
$$w^T = [1, 1, 1]$$



Exemplo

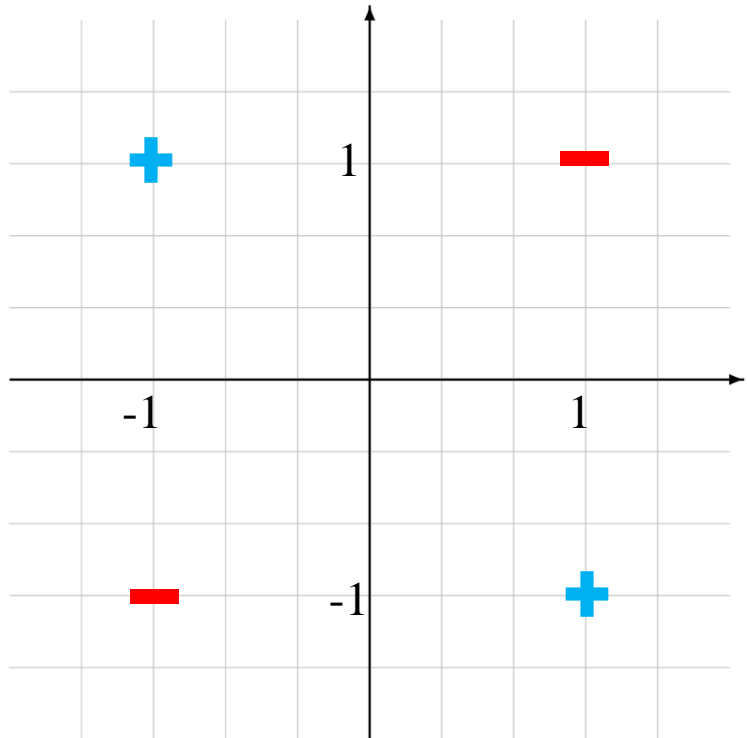
- Inicializar os pesos w aleatoriamente
- Enquanto existirem observações sendo classificadas de forma errada:
 - Selecionar observação $x^{(n)}$ tal que: $\text{sign}(x^{(n)}w) \neq y^{(n)}$
 - Atualizar os pesos: $w = w + y^{(n)}x^{(n)}$

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



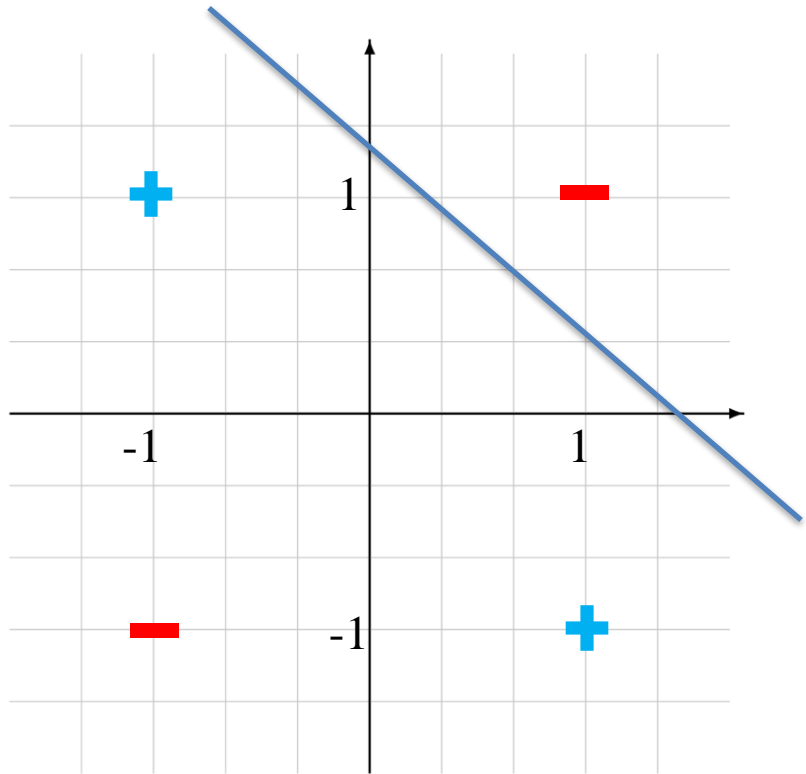
Limitação do Perceptron

- E se o nosso problema não for linearmente separável?
- XOR



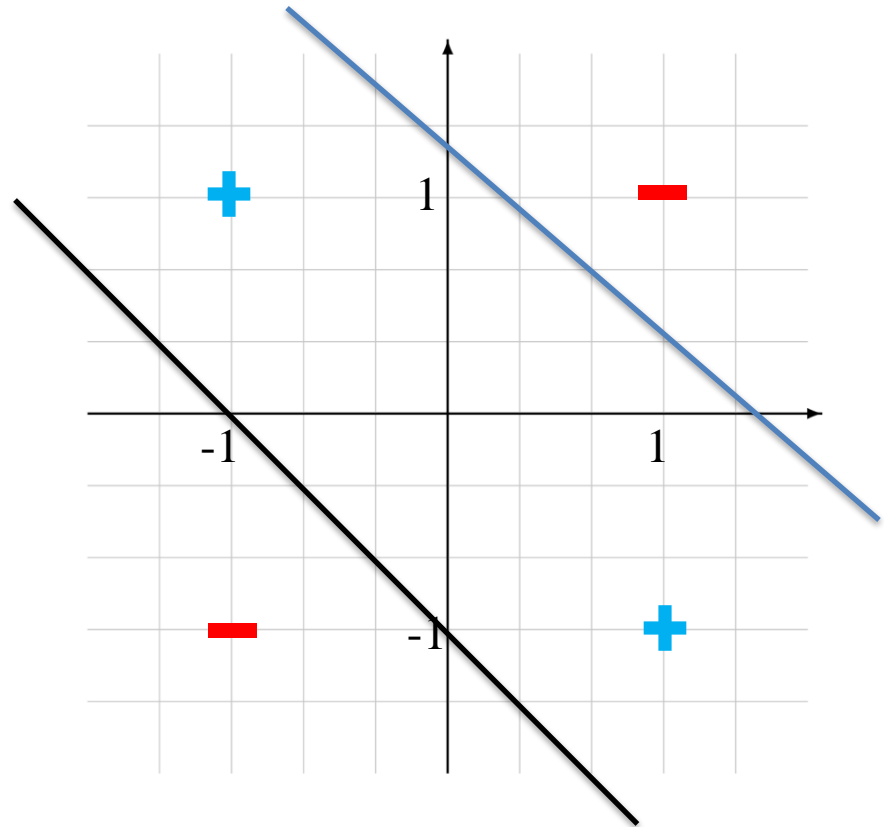
Limitação do Perceptron

- E se o nosso problema não for linearmente separável?
- XOR



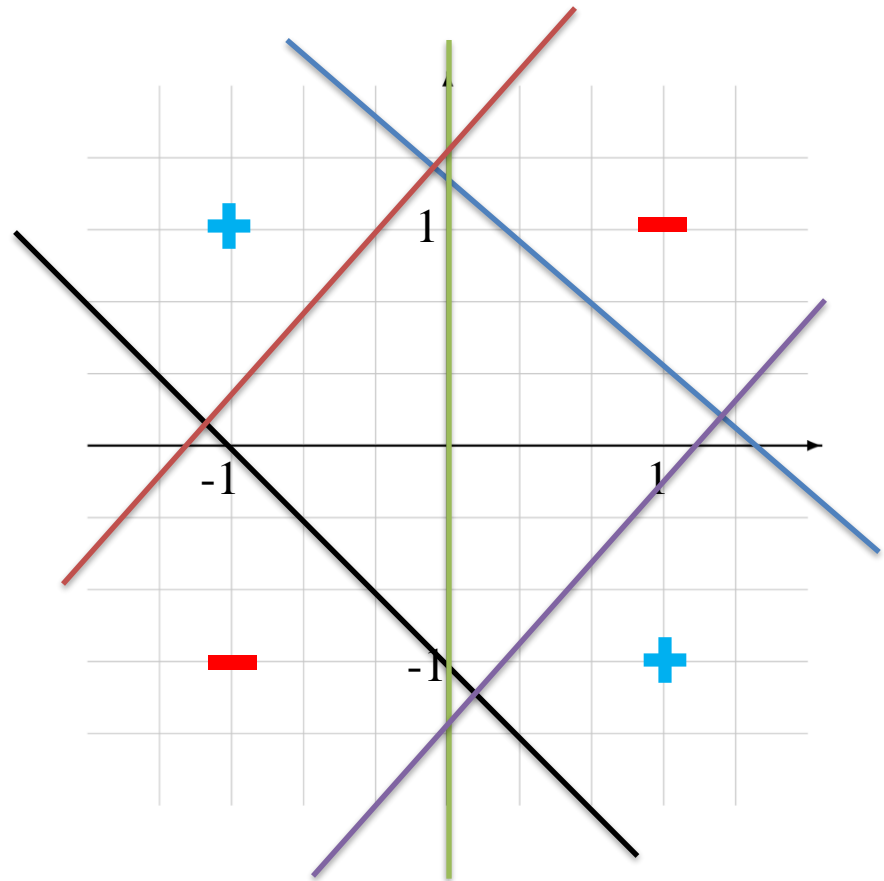
Limitação do Perceptron

- E se o nosso problema não for linearmente separável?
- XOR



Limitação do Perceptron

- E se o nosso problema não for linearmente separável?
- XOR

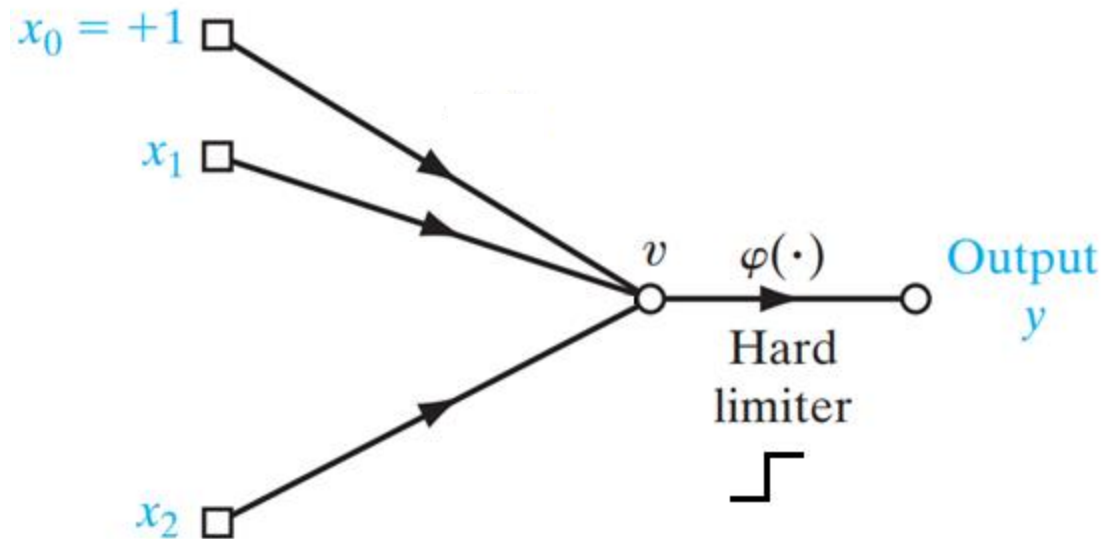


Como poderíamos resolver?

- XOR
 - Duas portas AND
 - Uma porta OR
- Conseguimos fazer isso com um Perceptron?

Como poderíamos resolver?

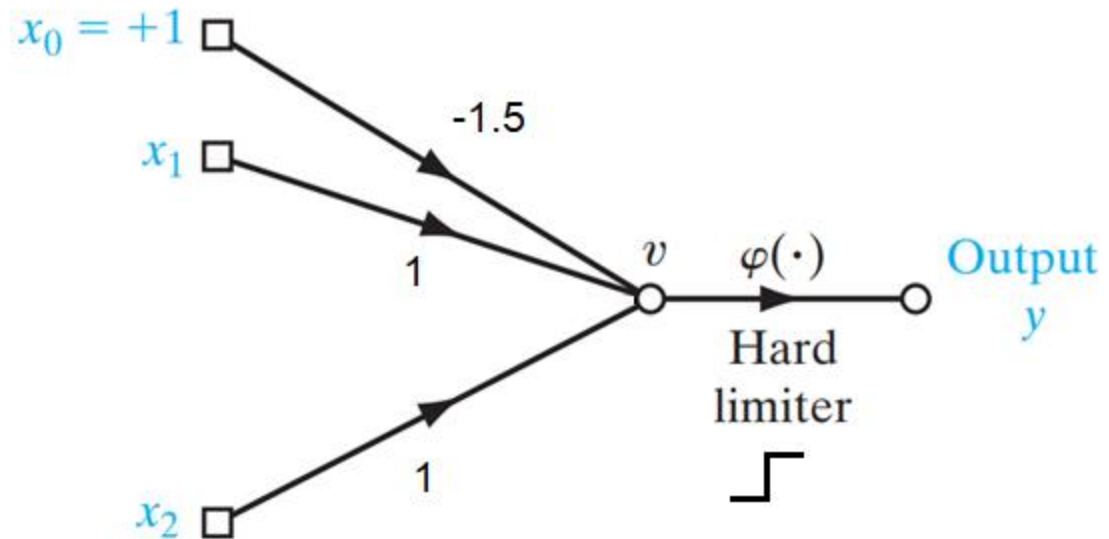
- Porta AND



x_1	x_2	y
-1	-1	0
-1	1	0
1	-1	0
1	1	1

Como poderíamos resolver?

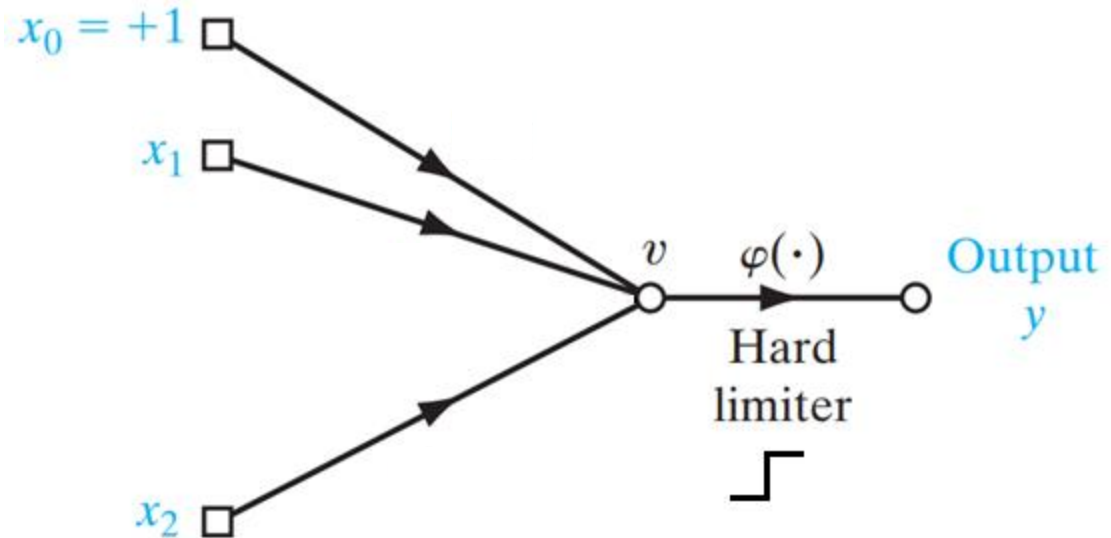
- Porta AND



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

Como poderíamos resolver?

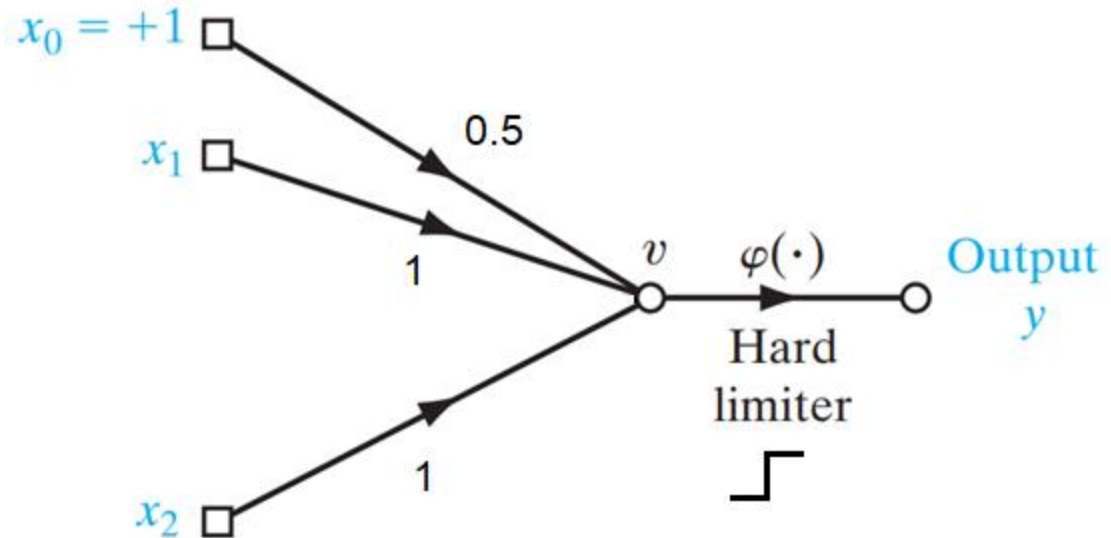
- Porta OR



x_1	x_2	y
-1	-1	0
-1	1	1
1	-1	1
1	1	1

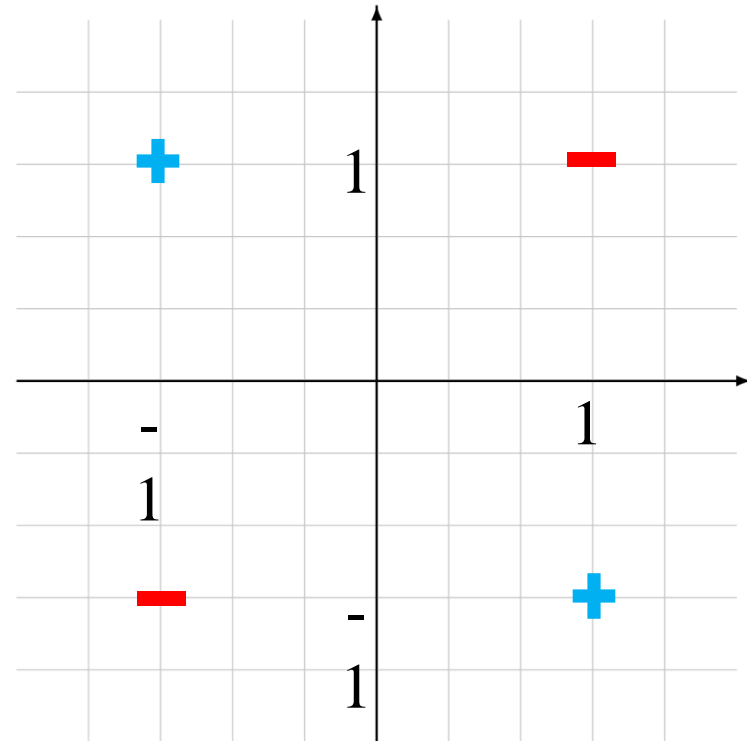
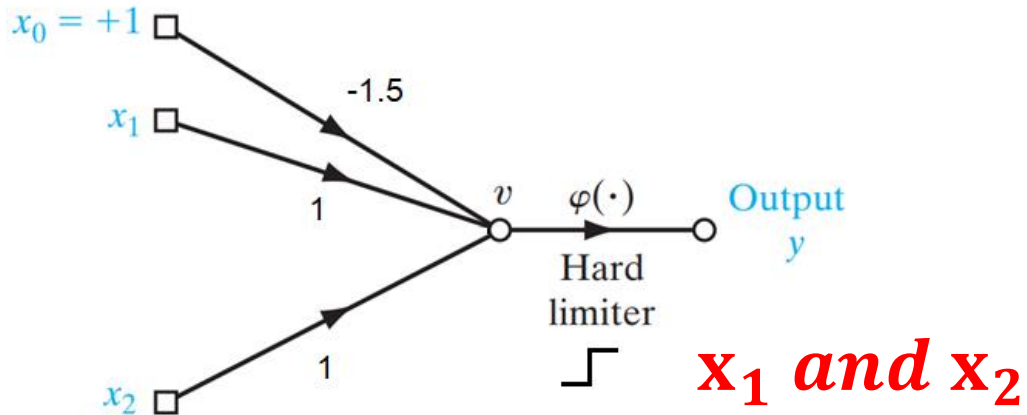
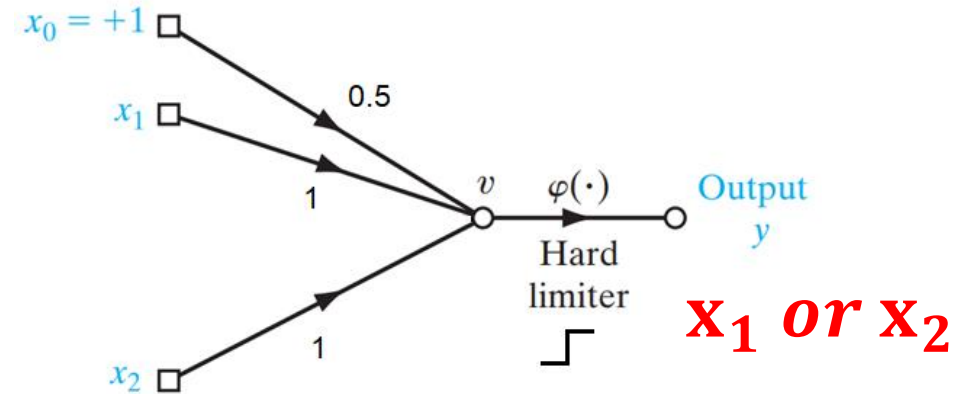
Como poderíamos resolver?

- Porta OR

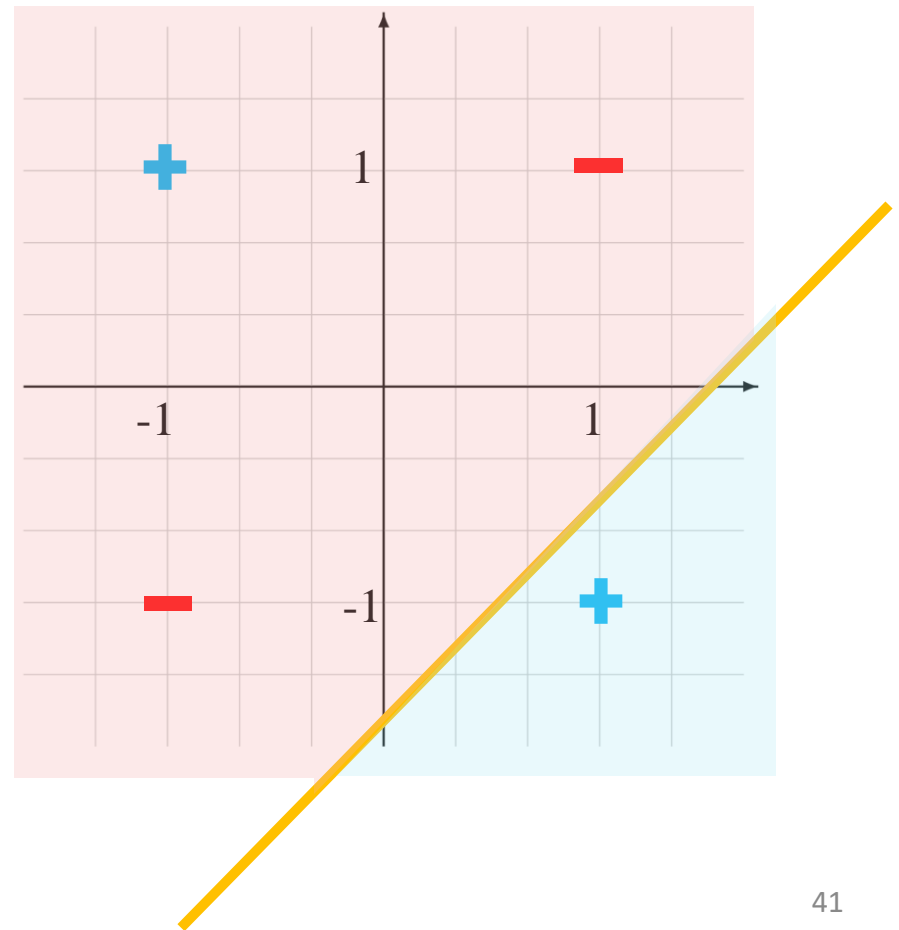
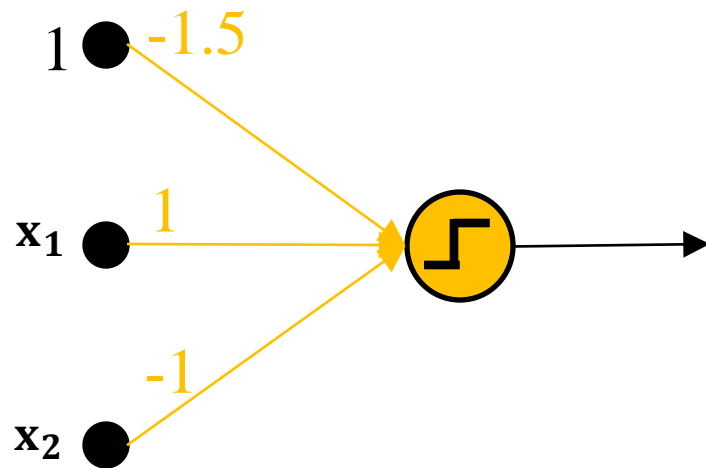


x_1	x_2	y
-1	-1	-1
-1	1	1
1	-1	1
1	1	1

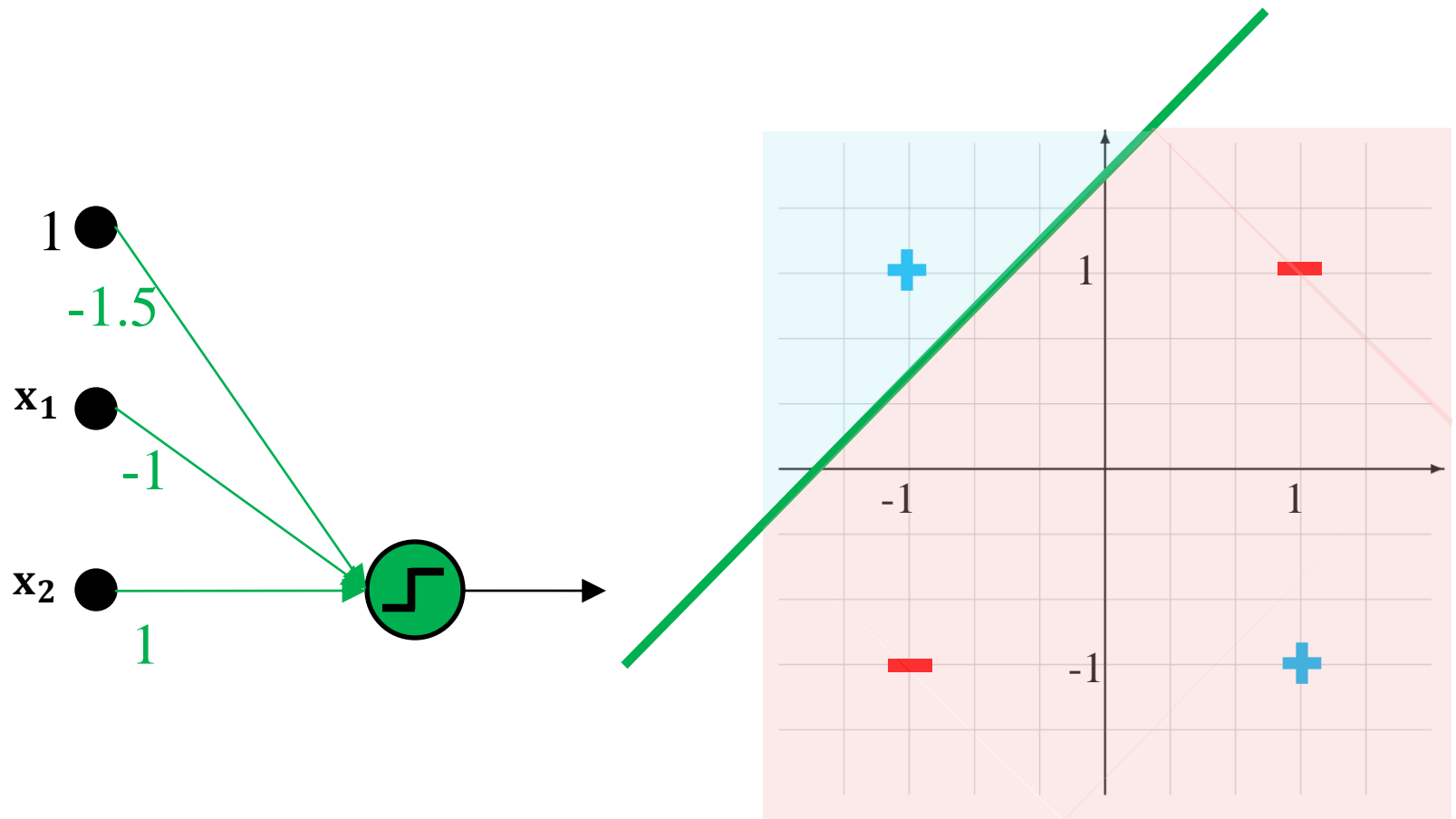
Como resolver



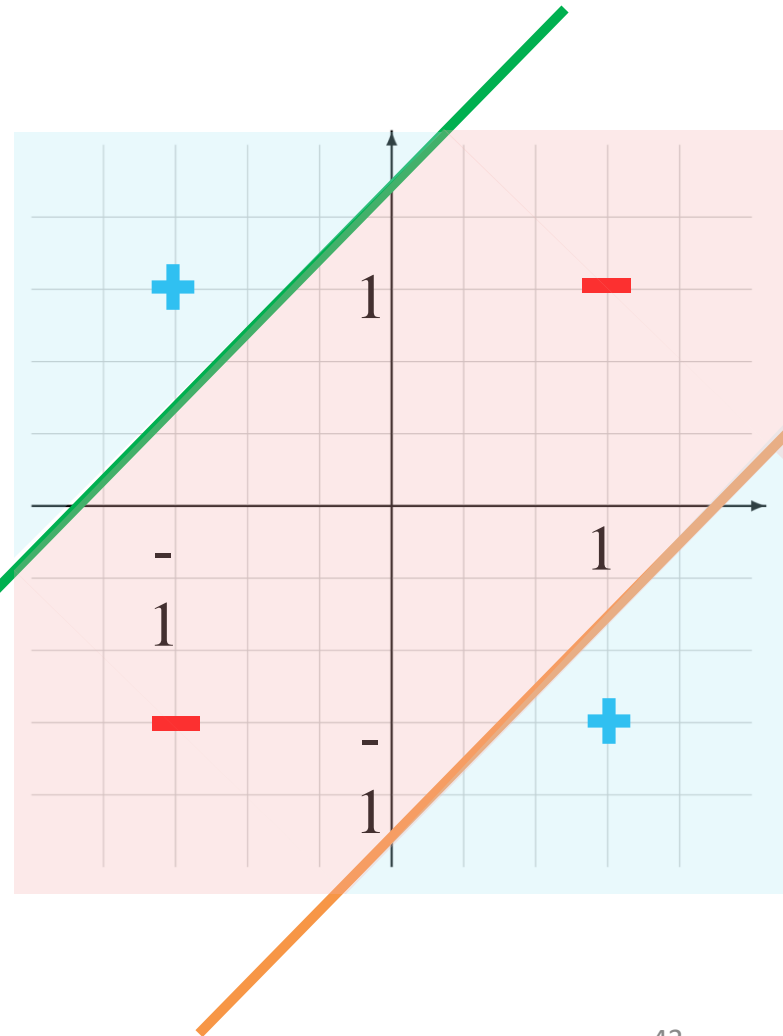
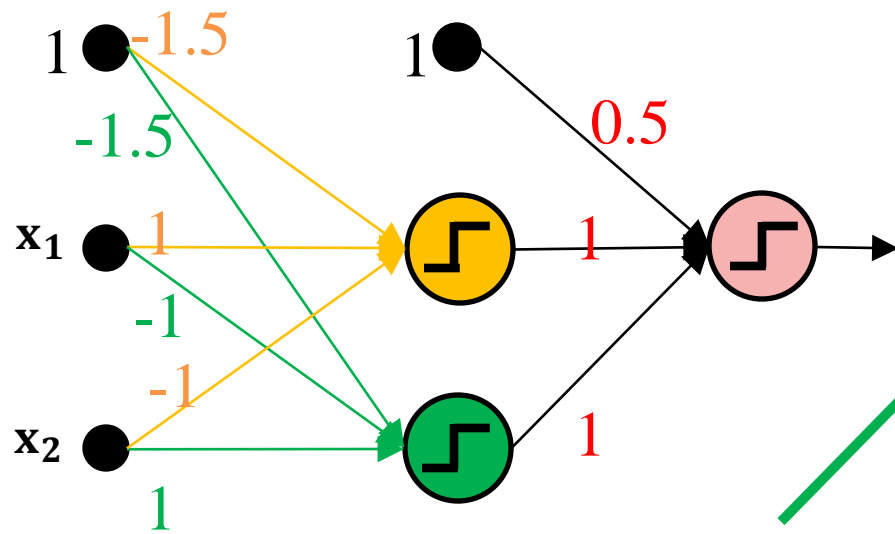
Como resolver



Como resolver

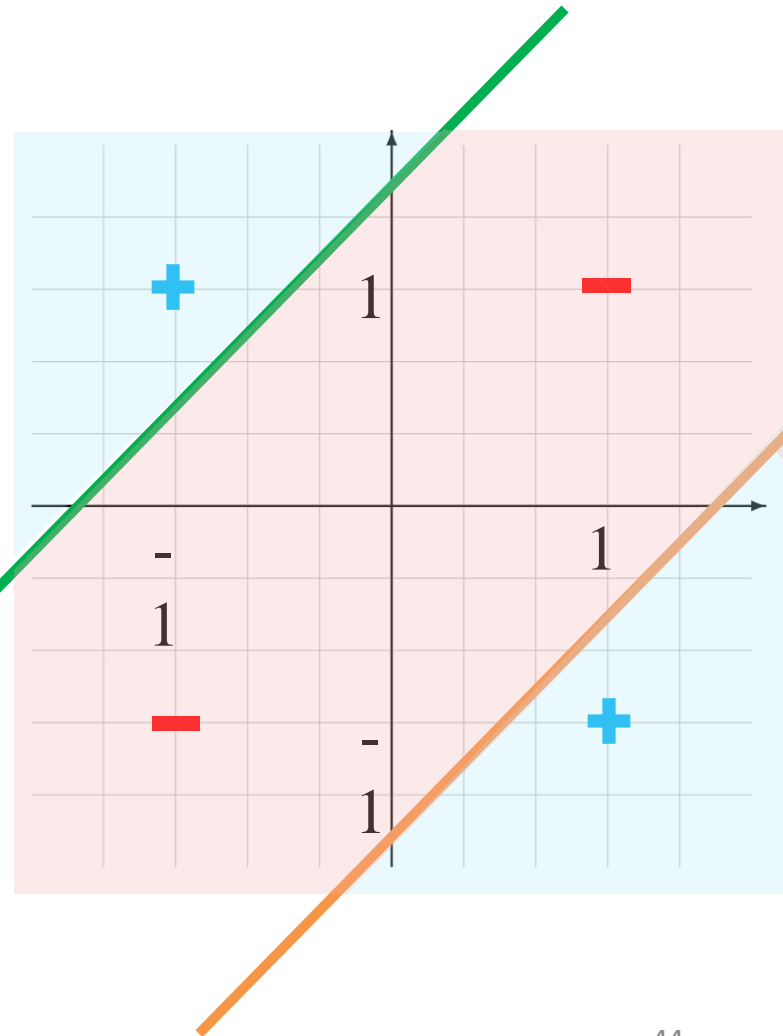
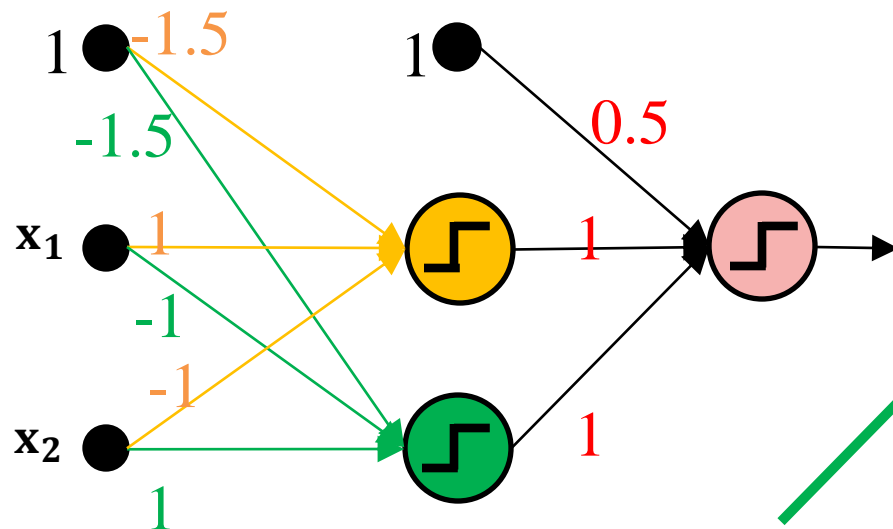


Como resolver



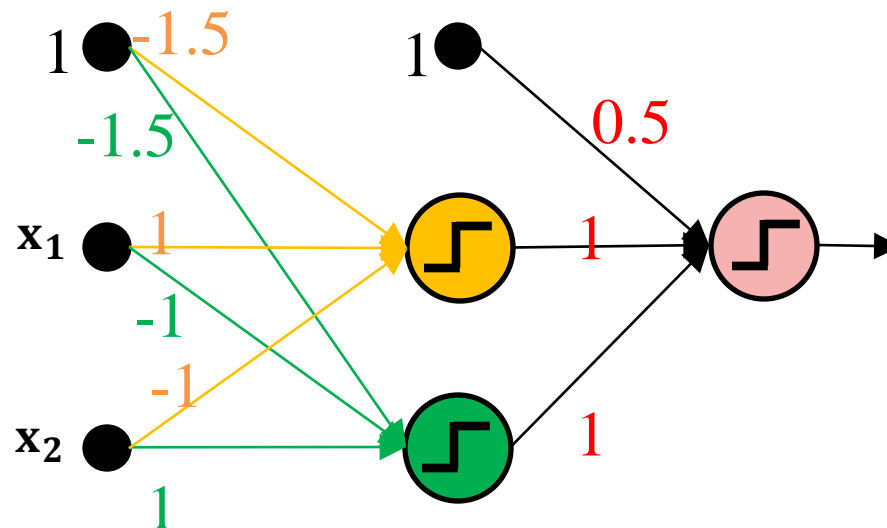
Como resolver

- Perfeito!



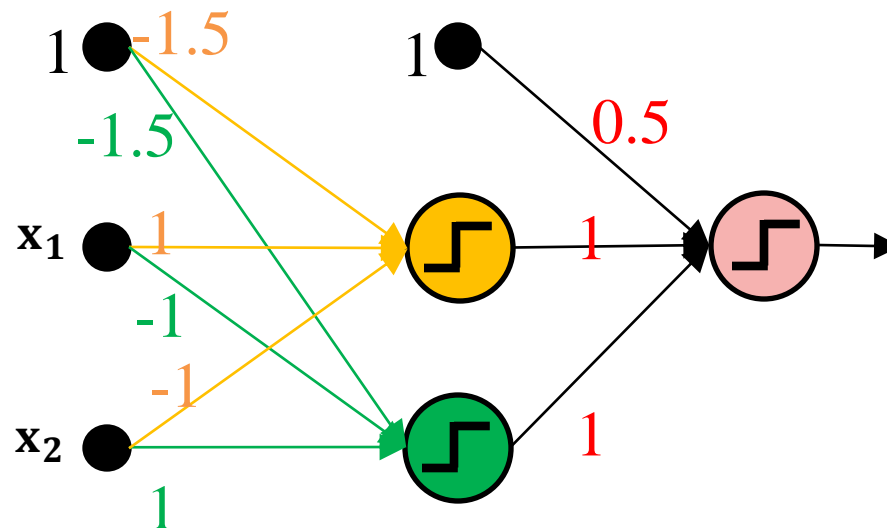
Perceptron Multi-Camada

- Agora conseguimos aproximar problemas não lineares.



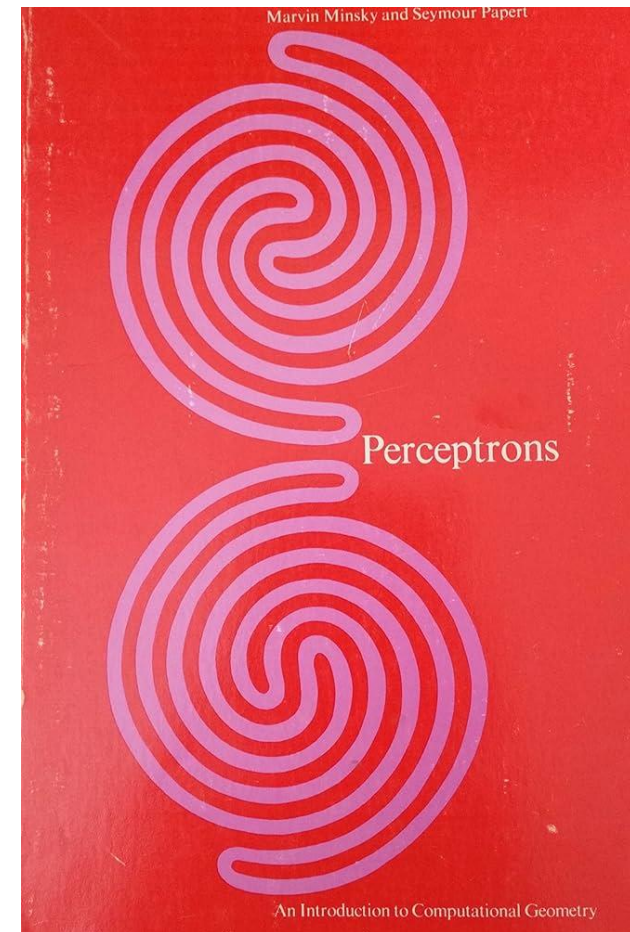
Perceptron Multi-Camada

- O problema é: como treinamos esse modelo?



Inverno de IA

- Como treinar uma MLP é algo difícil
- Livro de Marvin Minsky e Seymour Papert
 - Apontava as limitações do Perceptron
 - Traz o conceito de função de ativação

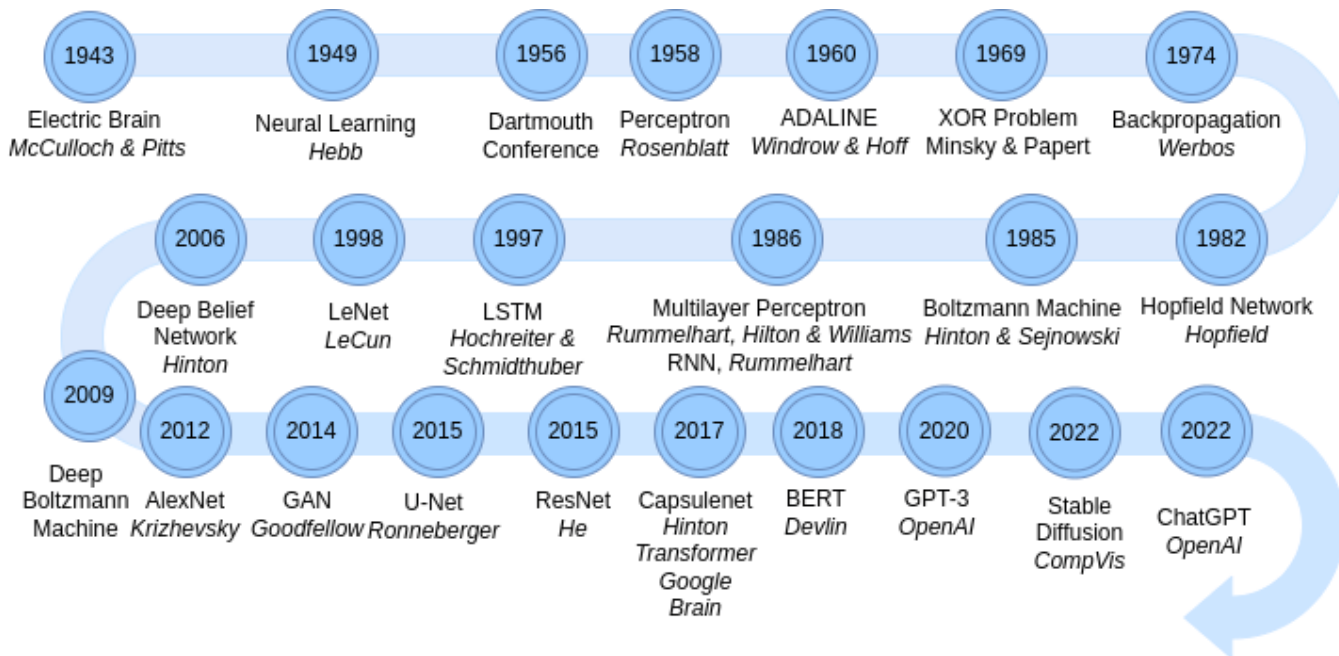


Inverno de IA

- Década de 70 foi marcada como o primeiro inverno de IA
 - 1974 – 1980
- Limitações do Perceptron
- Limitações de outras áreas de ML:
 - Falta de avanço em tradução de máquina
 - Falta de avanço em reconhecimento / compreensão de fala

Inverno de IA

- De acordo com Minsky e Papert Perceptrons Multi-camadas nunca seriam capazes de replicar a inteligência humana...



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.
- Slides adaptados dos originais dos profs. André Carvalho (ICMC-USP), Ricardo Campello (ICMC-USP), Andrew Ng (Stanford), Rodrigo C. Barros (PUCRS) e Lucas S. Kupssinskü (PUCRS)