

Treinamento – Classificação Linear

$$S = w_1 * x_1 + w_2 * x_2 - \theta$$

← Função de Saída

$$F = 1 \text{ se } s > 0, -1 \text{ se } s \leq 0$$

← Função de Ativação

$$\theta_i = \theta_i + \gamma(d - y)x_0/2$$

← Ajuste do Fator de Tendência

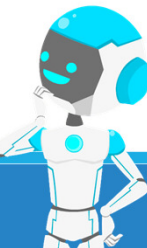
$$w_i = w_i + \gamma(d - y)x_i/2$$

← Ajuste dos Pesos

Y = saída

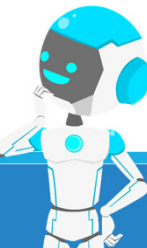
d = saída esperada

γ = taxa de aprendizado (próximo slide)



Taxa de Aprendizado (Step Size)

- Igual a Zero = Nunca vai convergir
- Muito baixo = algoritmo lento, pois vai fazer correções muito pequenas nos pesos
- Muito alto = algoritmo lento, pois vai fazer correções muito grandes nos pesos, ultrapassar o ideal



Classificação Linear - Binário

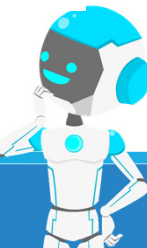
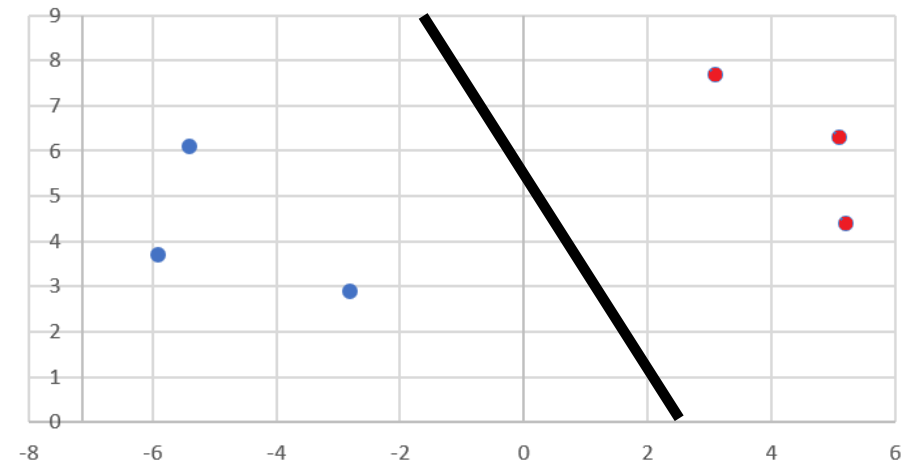


Orgânica



Não-orgânica

AtributoA	AtributoB	Classe
-5,4	6,1	0
-5,9	3,7	0
-2,8	2,9	0
3,1	7,7	1
5,1	6,3	1
5,2	4,4	1



Classificação Linear - Binário



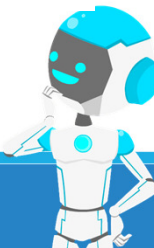
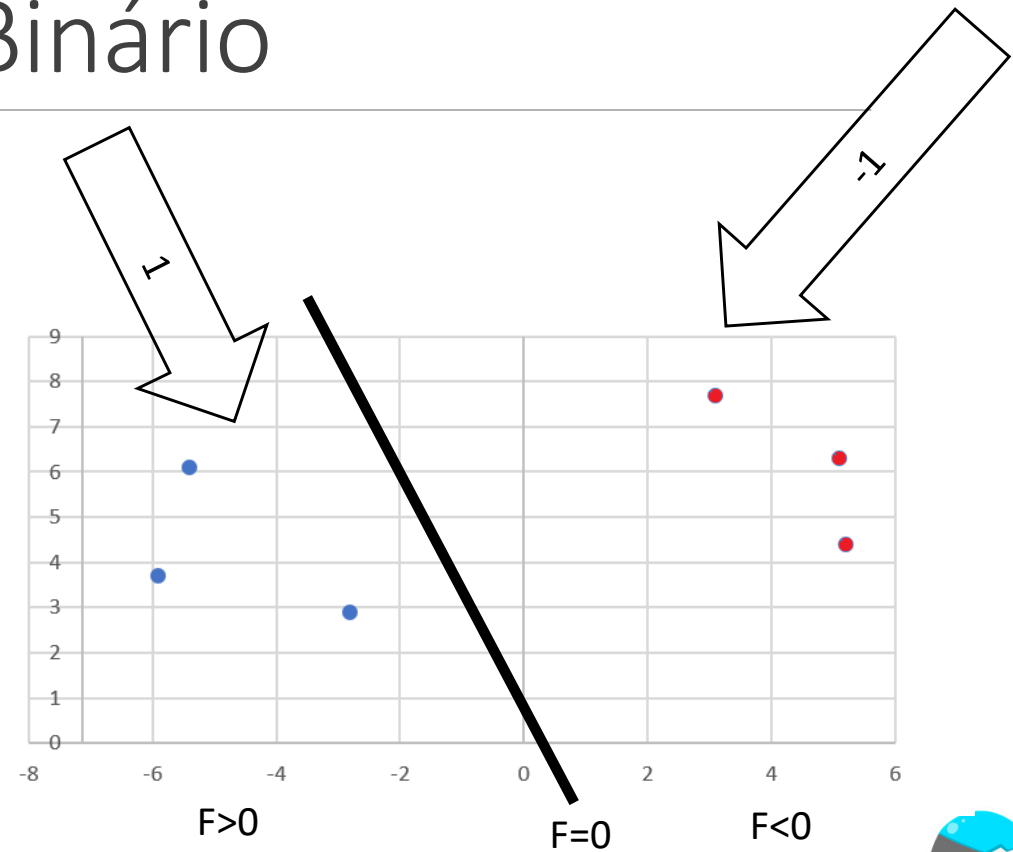
Não-orgânica



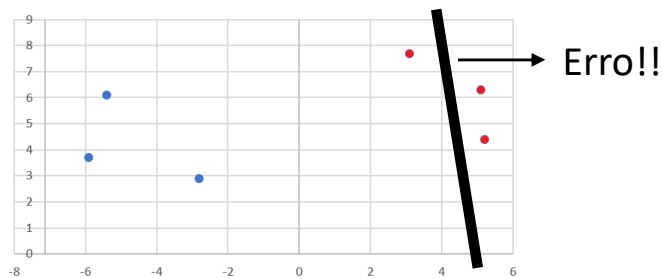
Orgânica

$$F = 1 \text{ se } s > 0, -1 \text{ se } s \leq 0$$

AtributoA	AtributoB	Classe
-5,4	6,1	0
-5,9	3,7	0
-2,8	2,9	0
3,1	7,7	1
5,1	6,3	1
5,2	4,4	1

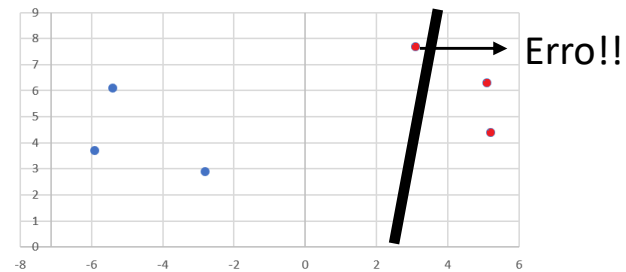


Ajustes dos Pesos



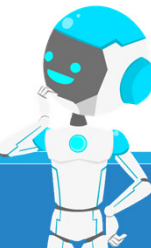
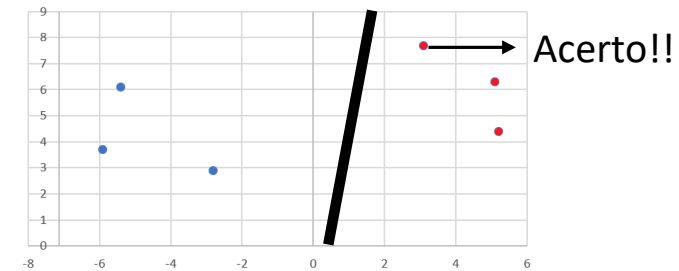
$$\theta_i = \theta_i + \gamma(d - y)x_0/2$$

$$w_i = w_i + \gamma(d - y)x_i/2$$

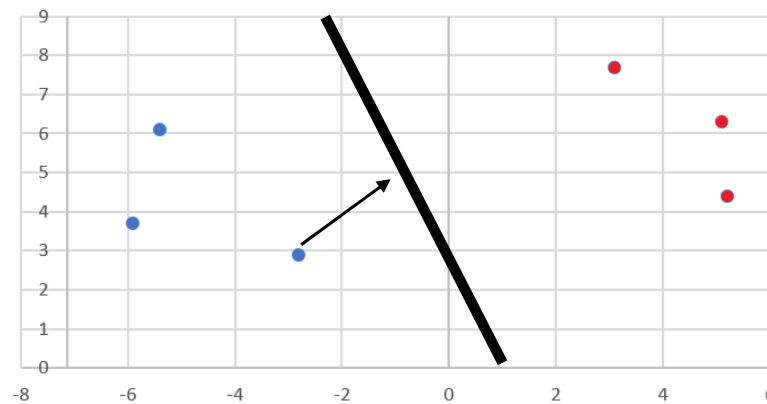


$$\theta_i = \theta_i + \gamma(d - y)x_0/2$$

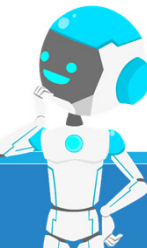
$$w_i = w_i + \gamma(d - y)x_i/2$$



Margem



Distância mínima entre uma instância e o hiperplano, quanto mais próximos, mas difícil é o treino do modelo



Aprendizado Online vs Batch

- Online: Os pesos (modelo) podem ser atualizados a cada instancia (Final da Iteração)
- Batch: Os pesos só são atualizados após todas as instâncias (Final da Epoch)

