

# Máquinas de vetores de suporte (SVM)

Jones Granatyr

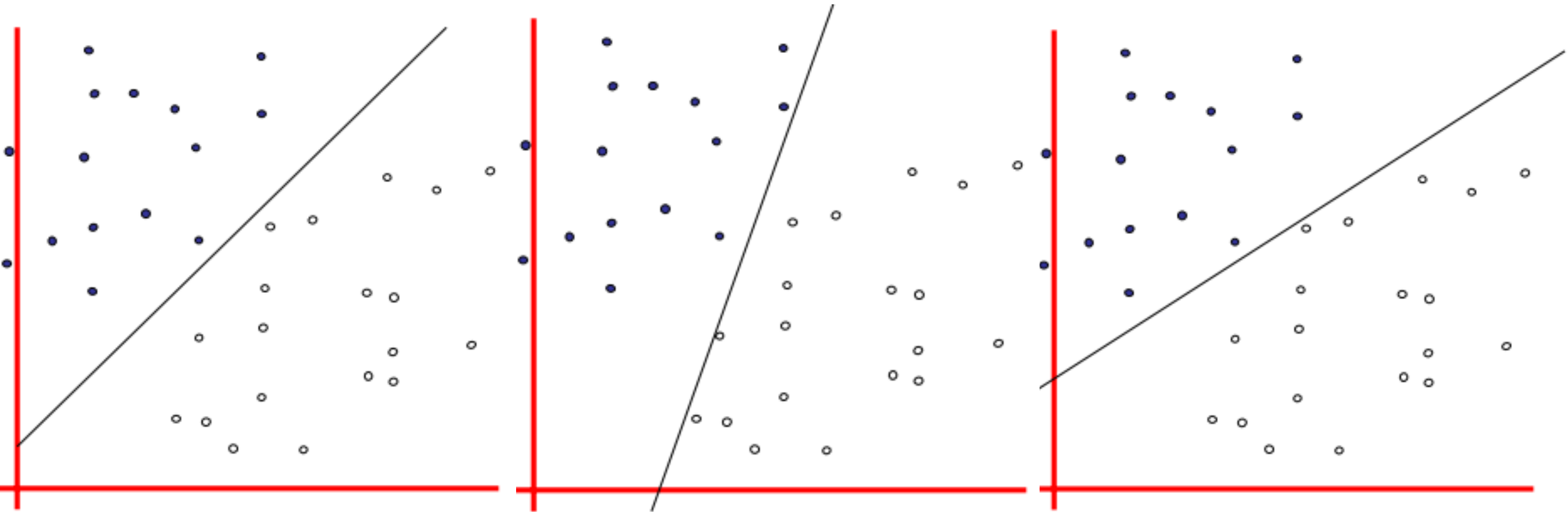


# SVM – introdução

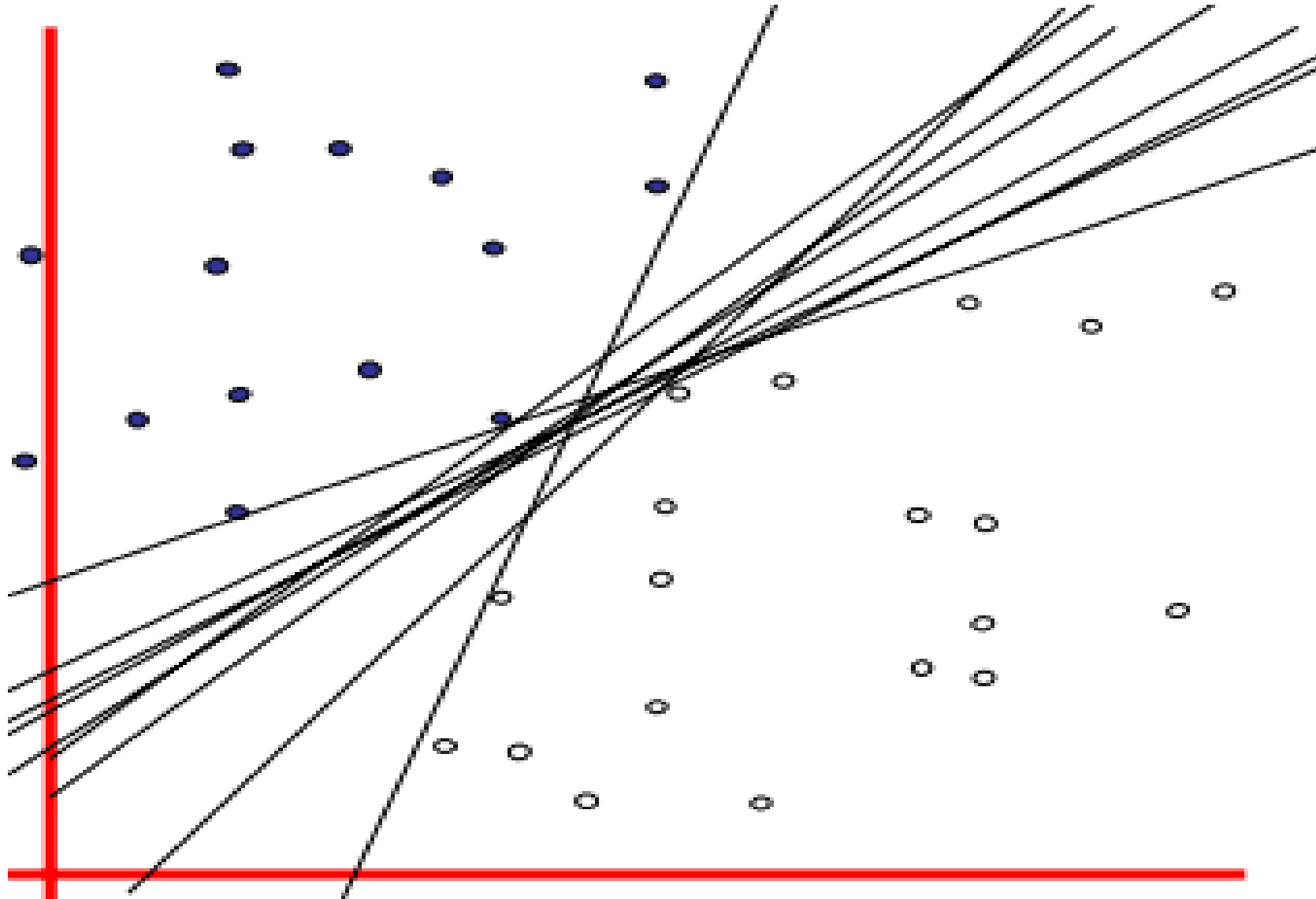
- Em geral supera outros algoritmos de aprendizagem de máquina
- Tarefas complexas: reconhecimento de caracteres, voz, imagens
- Considerado por vários anos como o mais eficiente algoritmo

# SVM – introdução

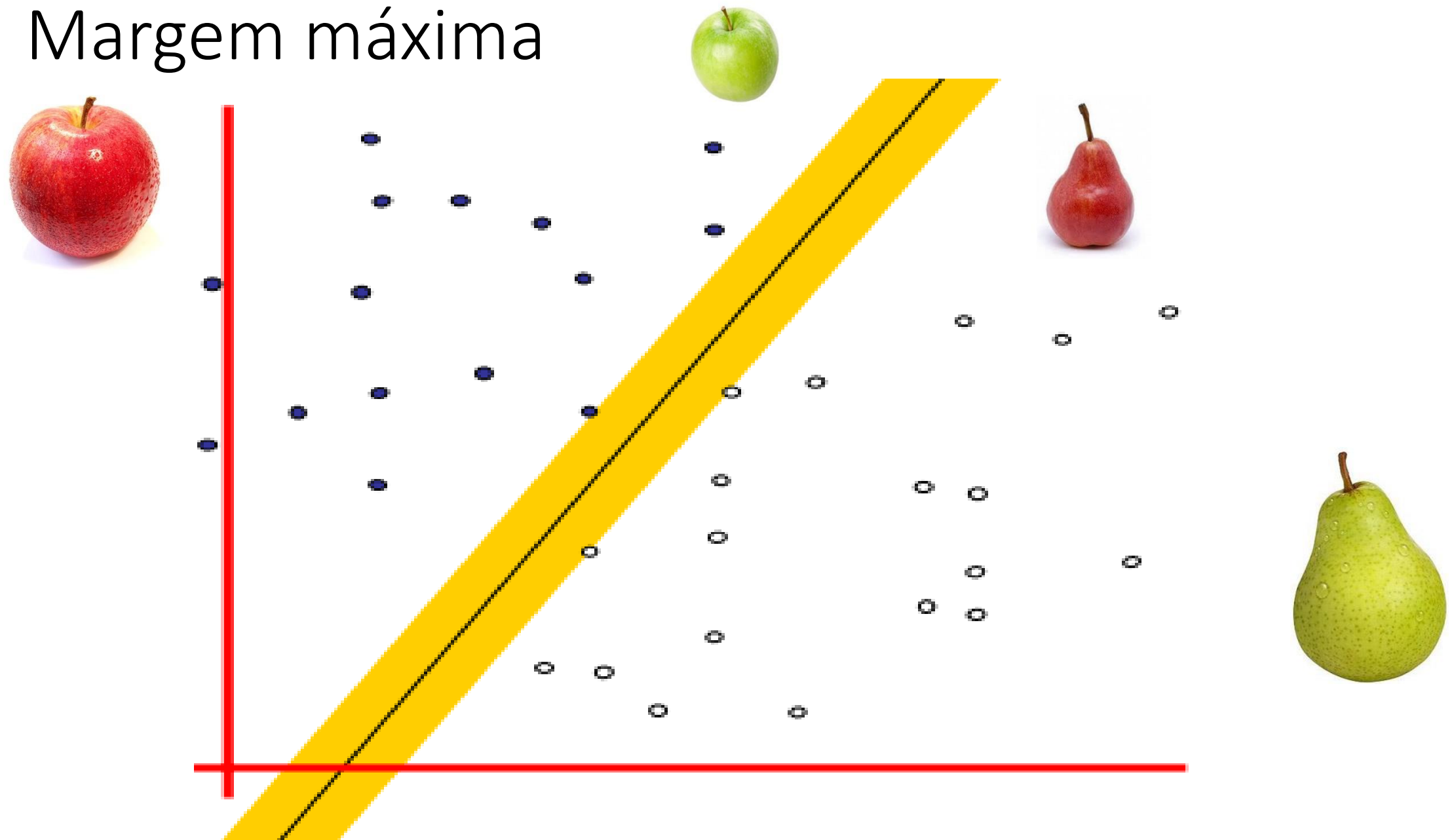
- Aprende hiperplanos de separação com margem máxima



Qual o melhor hiperplano?

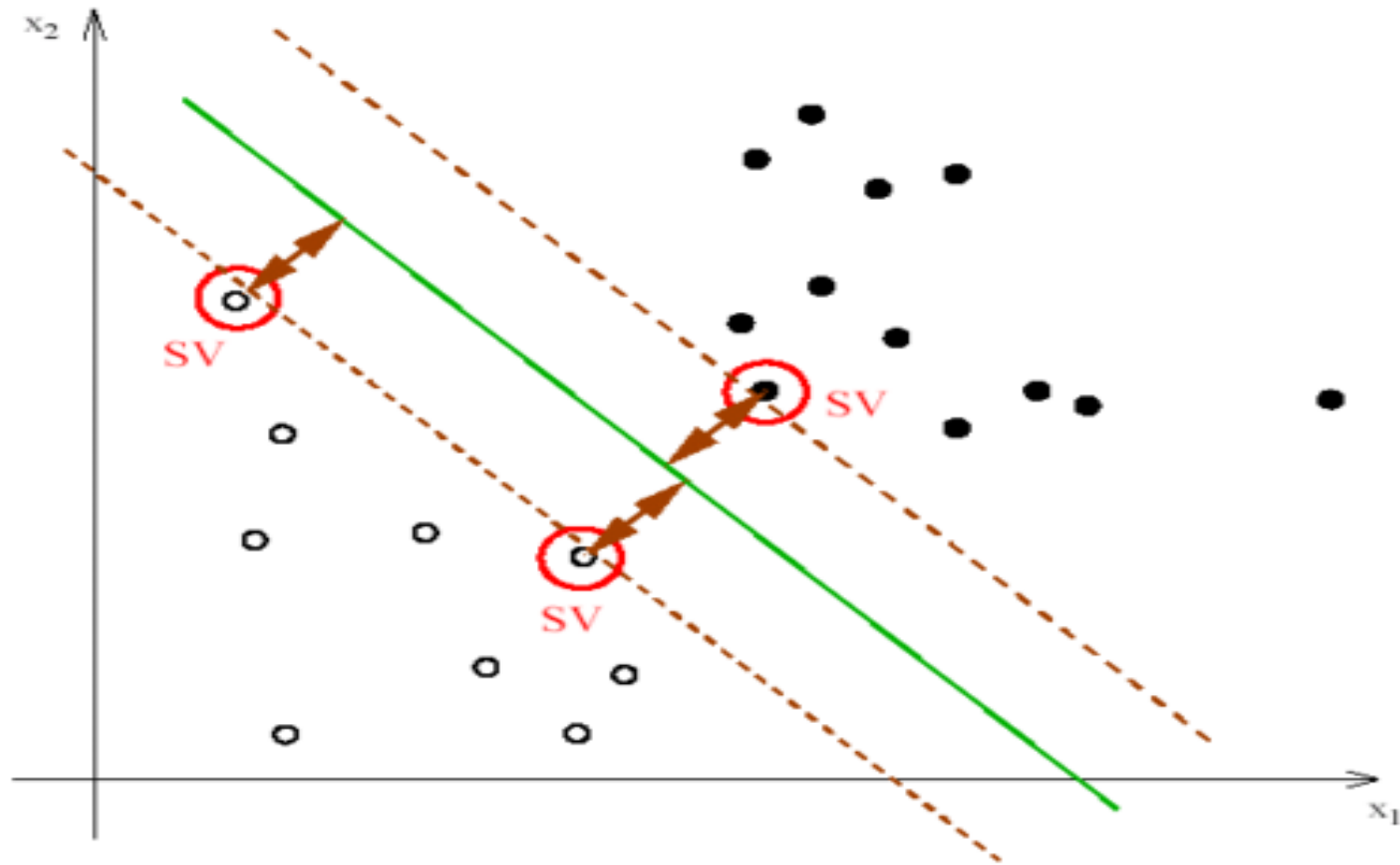


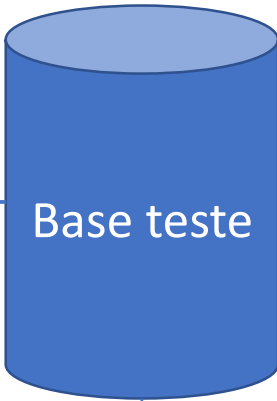
# Margem máxima



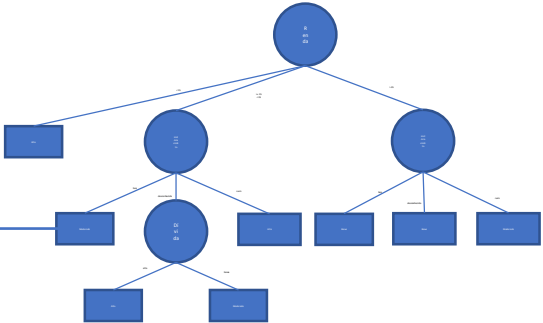
# Vetores de suporte

- Reconstrução do hiperplano

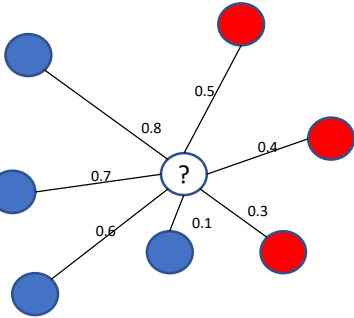




Risco de crédito	História do crédito			Dívida		Garantias		Renda anual		
	Boa	Desconhecida	Ruim	Até 2	Mais 2	Sei não	Até 1000	> 10000	> 10000	> 10000
Até 1/5	1/5	2/5	3/4	4/7	2/7	6/11	0	3/3	2/4	1/7
Entre 1/5 e 2/5	1/5	1/5	1/4	1/7	2/7	2/11	1/3	0	2/4	1/7
Entre 2/5 e 3/5	3/5	2/5	0	2/7	3/7	3/11	2/3	0	0	5/7



Regra	Resultado
Se renda = >35.000 E história_crédito = BOA	Risco = BAIXO
Se renda = >35.000 e história_crédito = DESCONHECIDA	Risco = BAIXO
Default (padrão)	Risco = ALTO



Naive bayes

Árvore de decisão

Regras

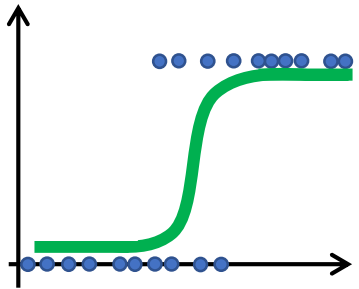
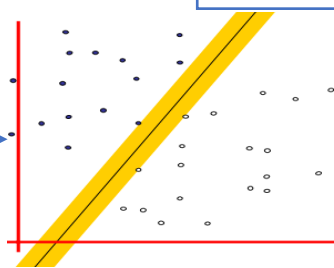
kNN

Regressão logística

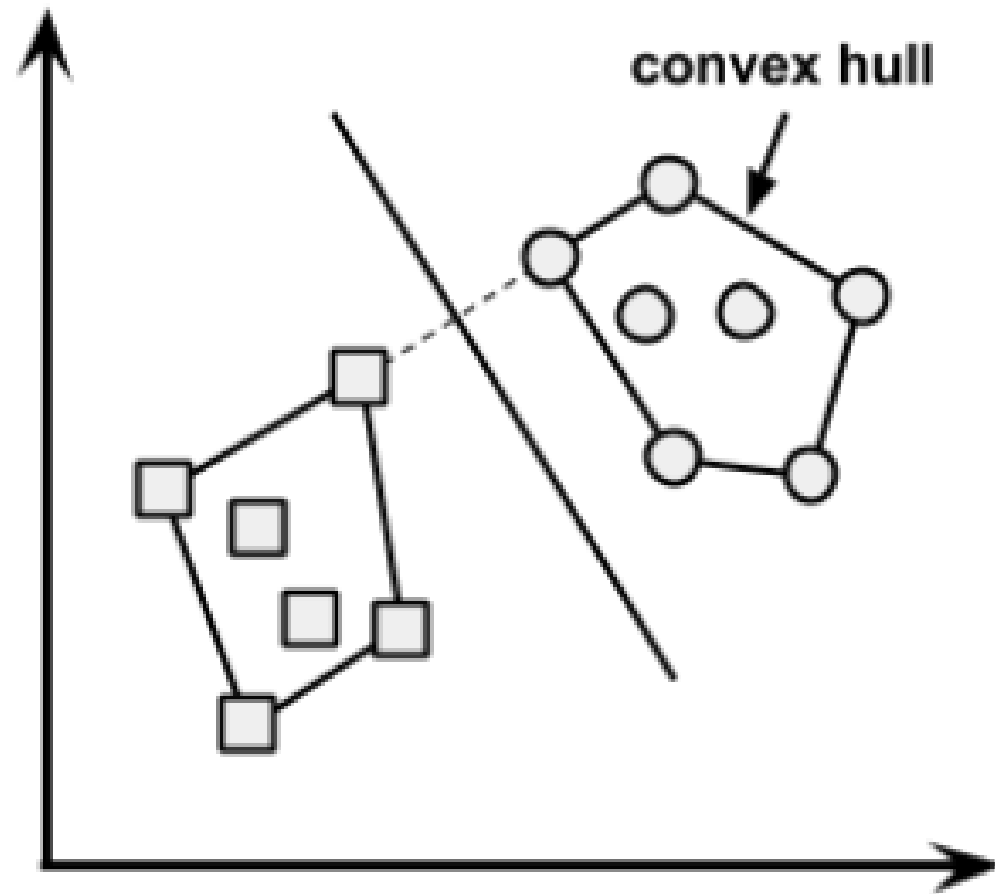
SVM

Registros  
% acerto

Base teste

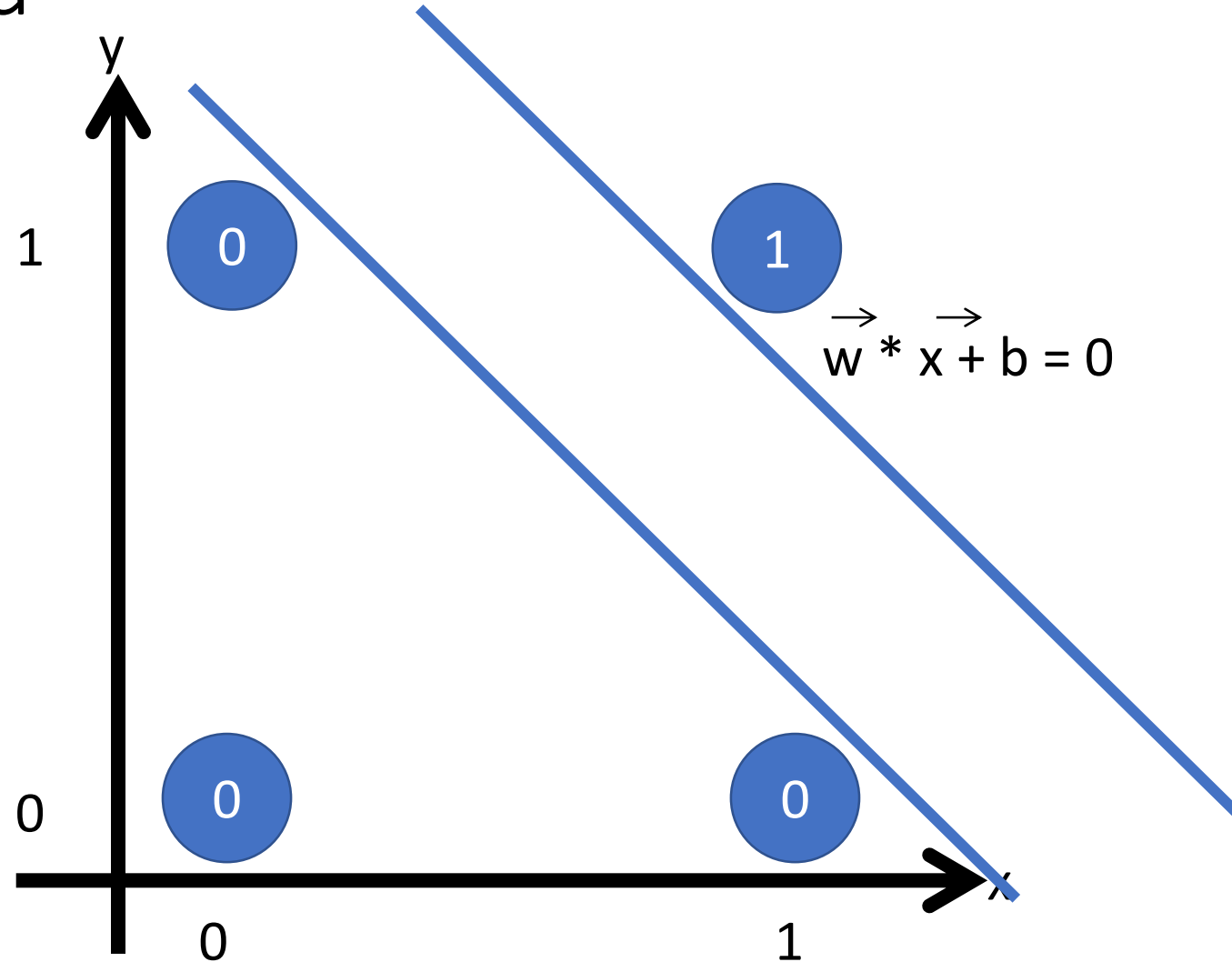


# Criação do hiperplano – Convex hulls (envoltória convexa)





# Criação do hiperplano – abordagem matemática



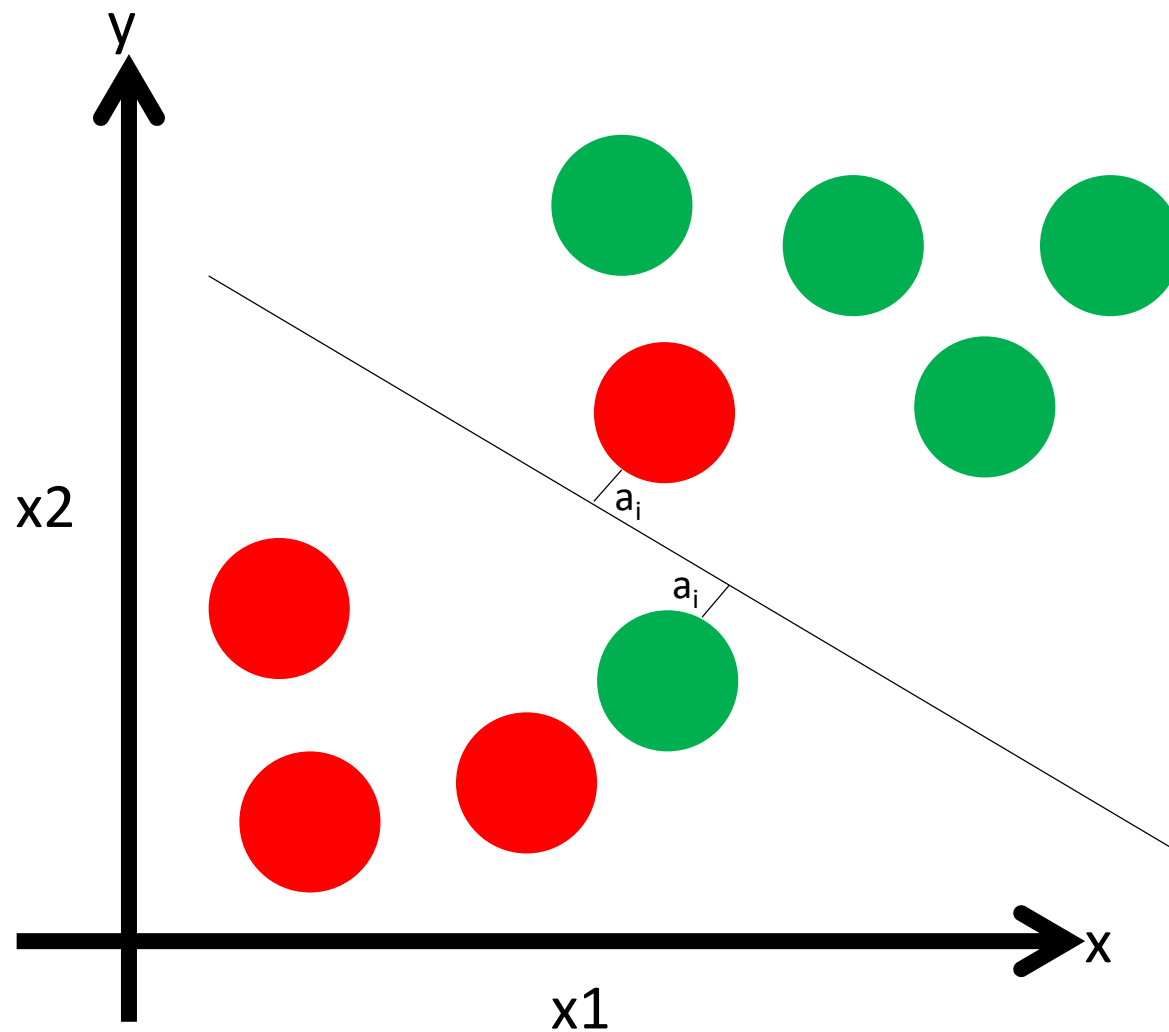
# Erros e custo

$$\frac{1}{2} |w|^2 + c \sum_i a_i$$

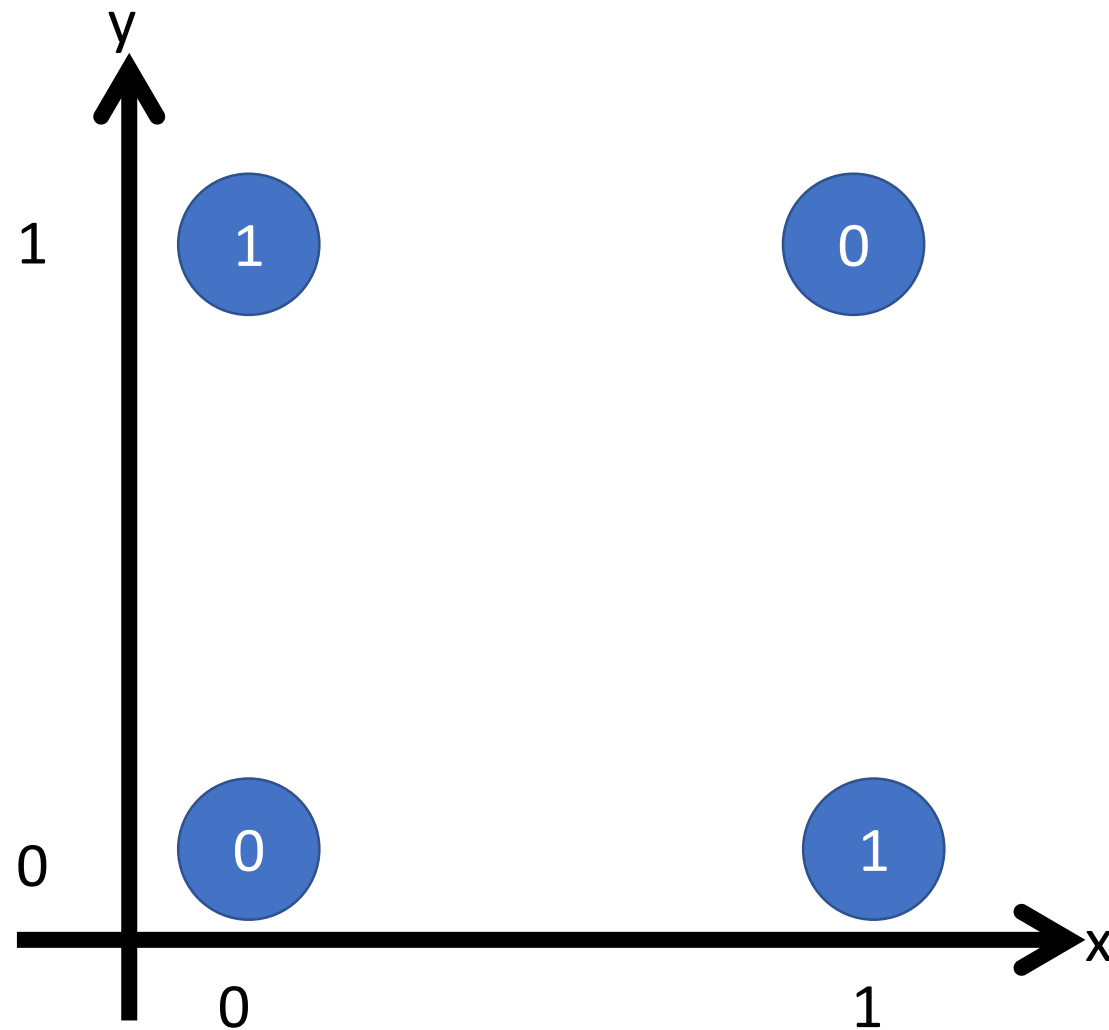
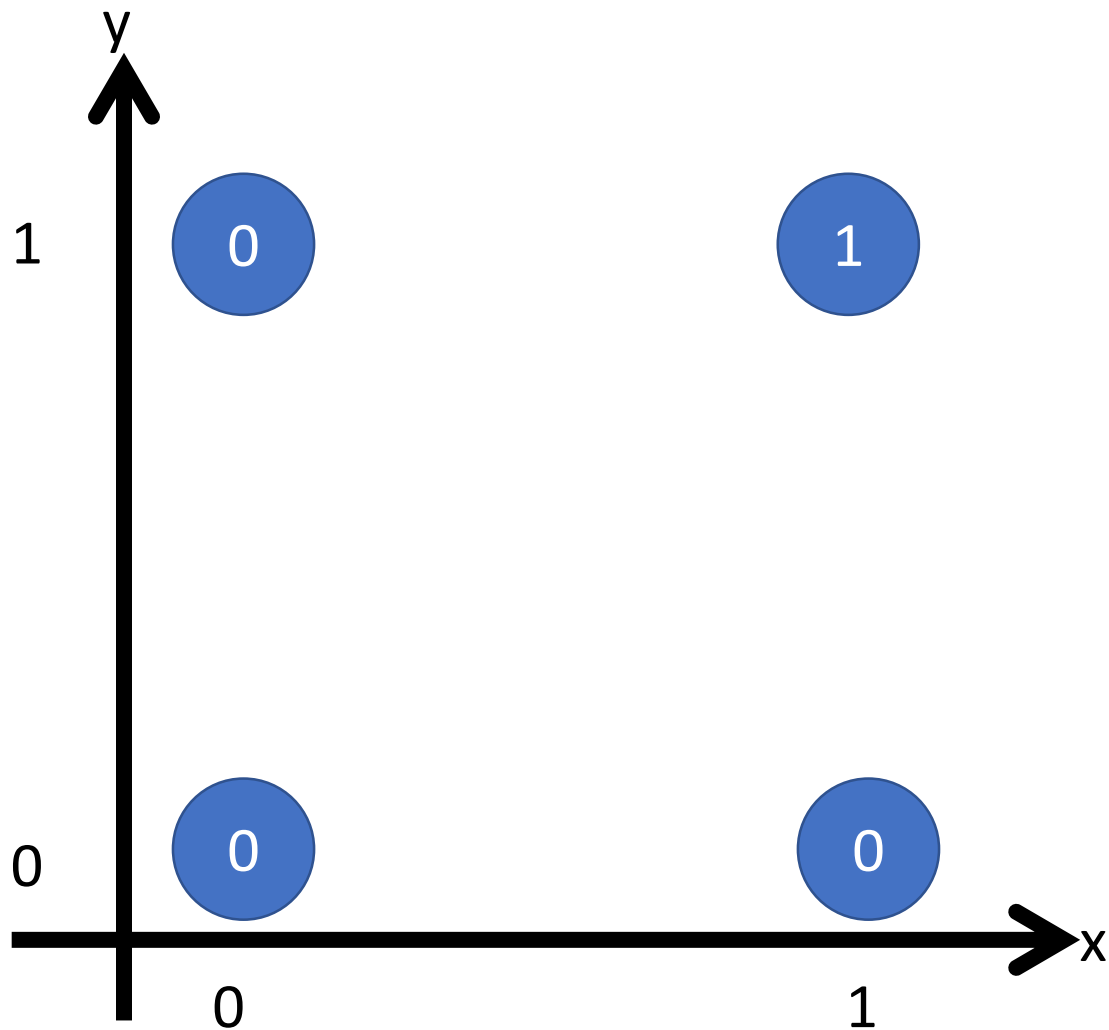
$c$  = punição por classificação incorreta

$c$  alto = tenta 100% de separação

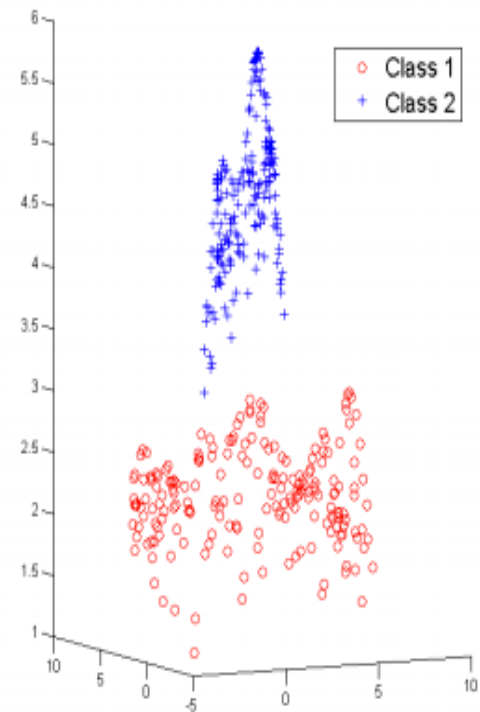
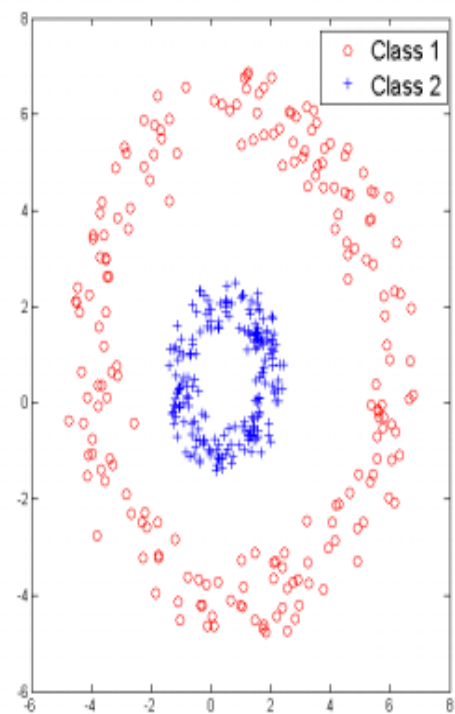
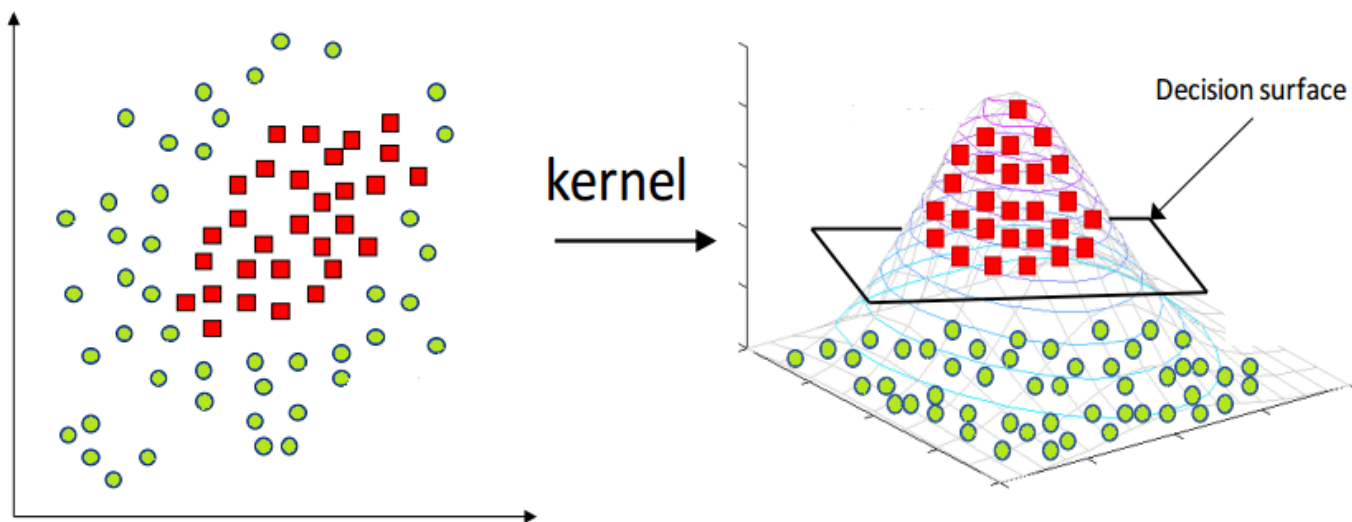
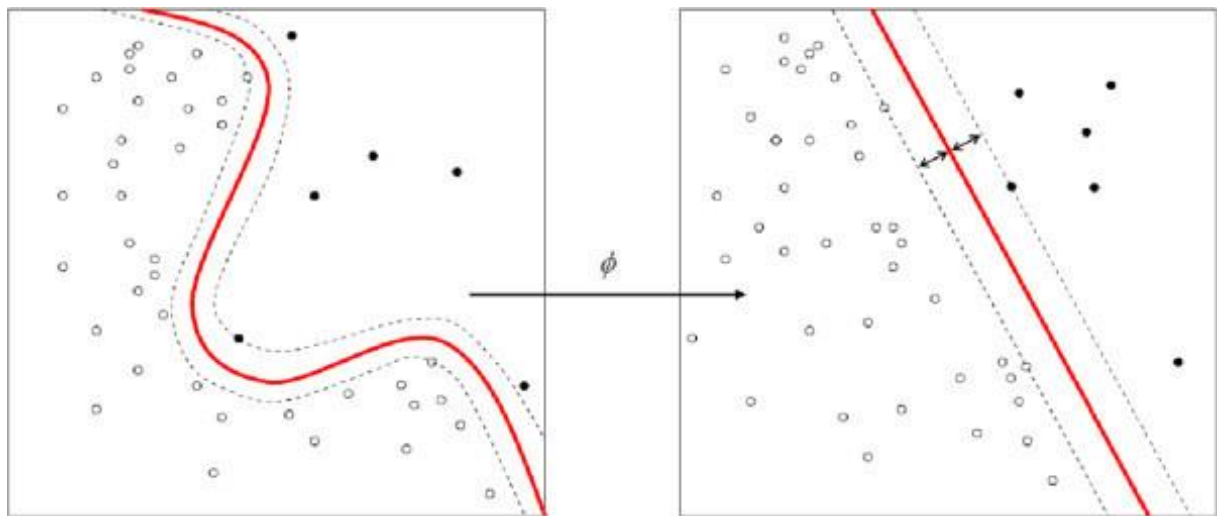
$c$  baixo = permite mais erros



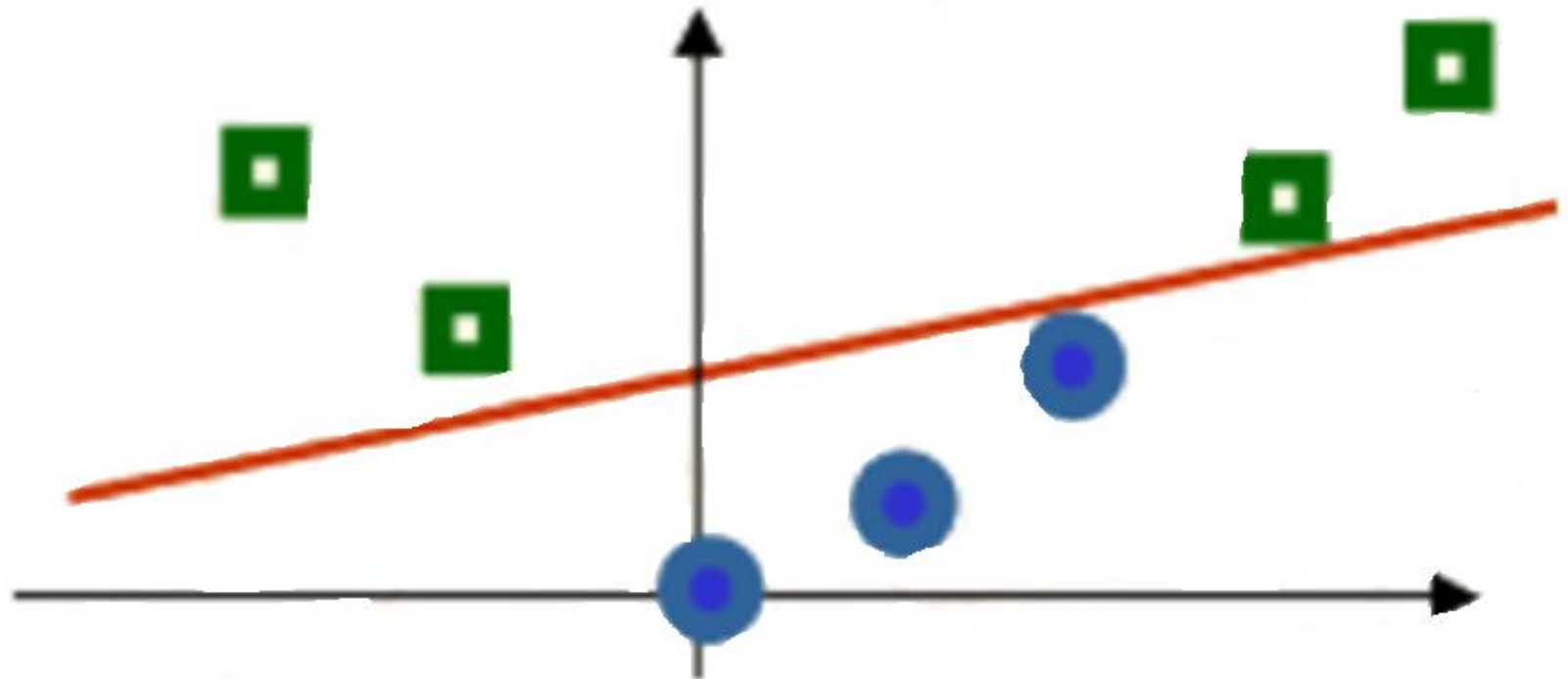
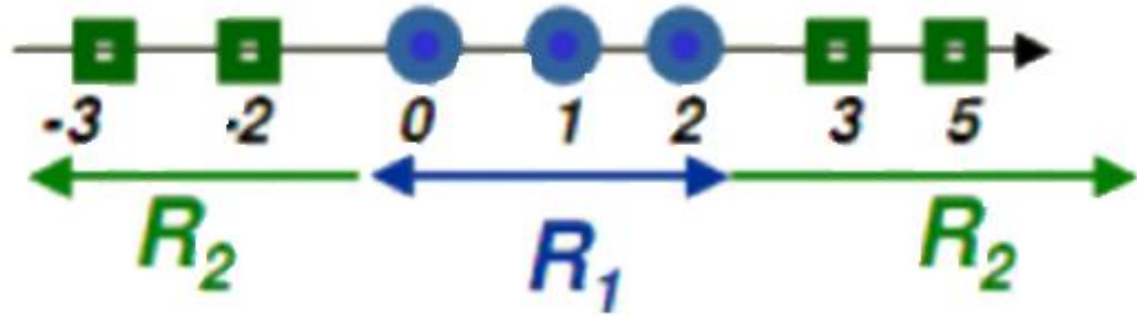
# Linear x Não linear



# SVMs não lineares (Kernel Trick)

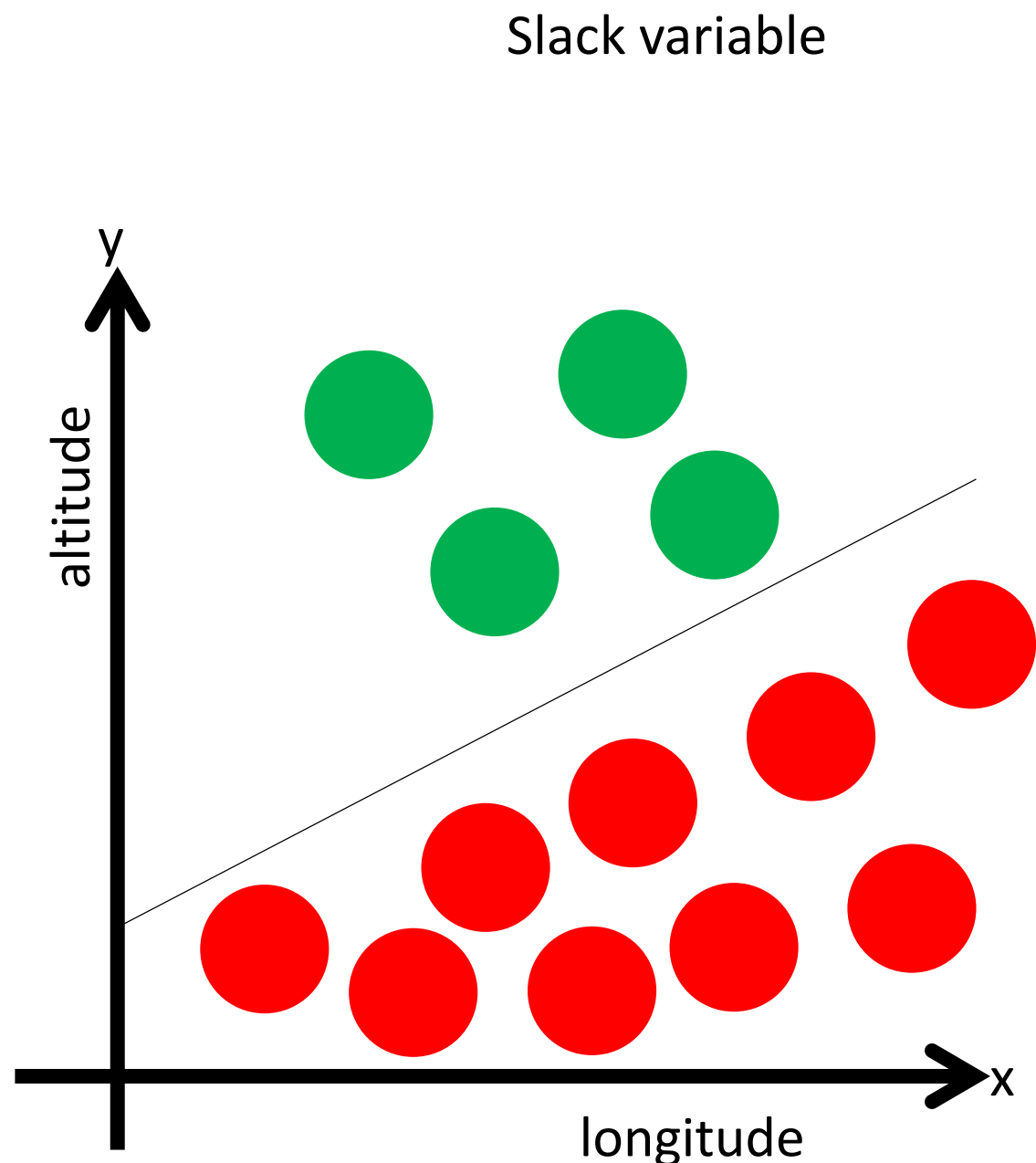
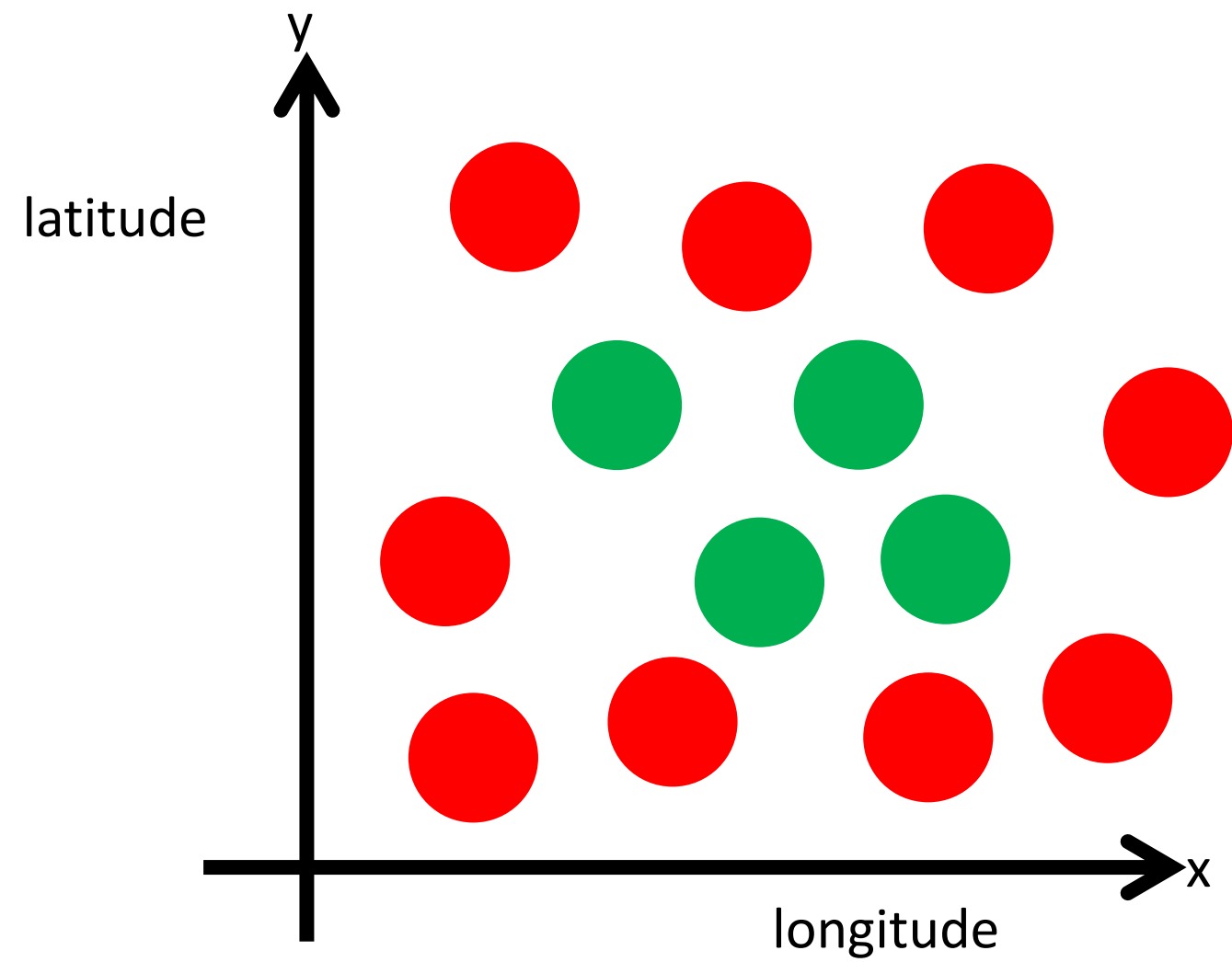


# SVMs não lineares (Kernel Trick)



# Kernels

Kernel	Inner Product Kernel
Linear	$K(x, y) = (x \cdot y)$
Gaussian	$K(x, y) = \exp\left(-\frac{\ x - x_i\ ^2}{2\sigma^2}\right)$
Polynomial	$K(x, y) = (x \cdot y)^p$
Tangent Hyperbolic	$K(x, y) = \tanh(x \cdot y - \Theta)$



# SVM

- Vantagens
  - Não é muito influenciado por ruídos nos dados
  - Utilizado para classificação e regressão
  - Aprende conceitos não presentes nos dados originais
  - Mais fácil de usar do que redes neurais
- Desvantagens
  - Testar várias combinações de parâmetros
  - Lento
  - Black box



# Conclusão

