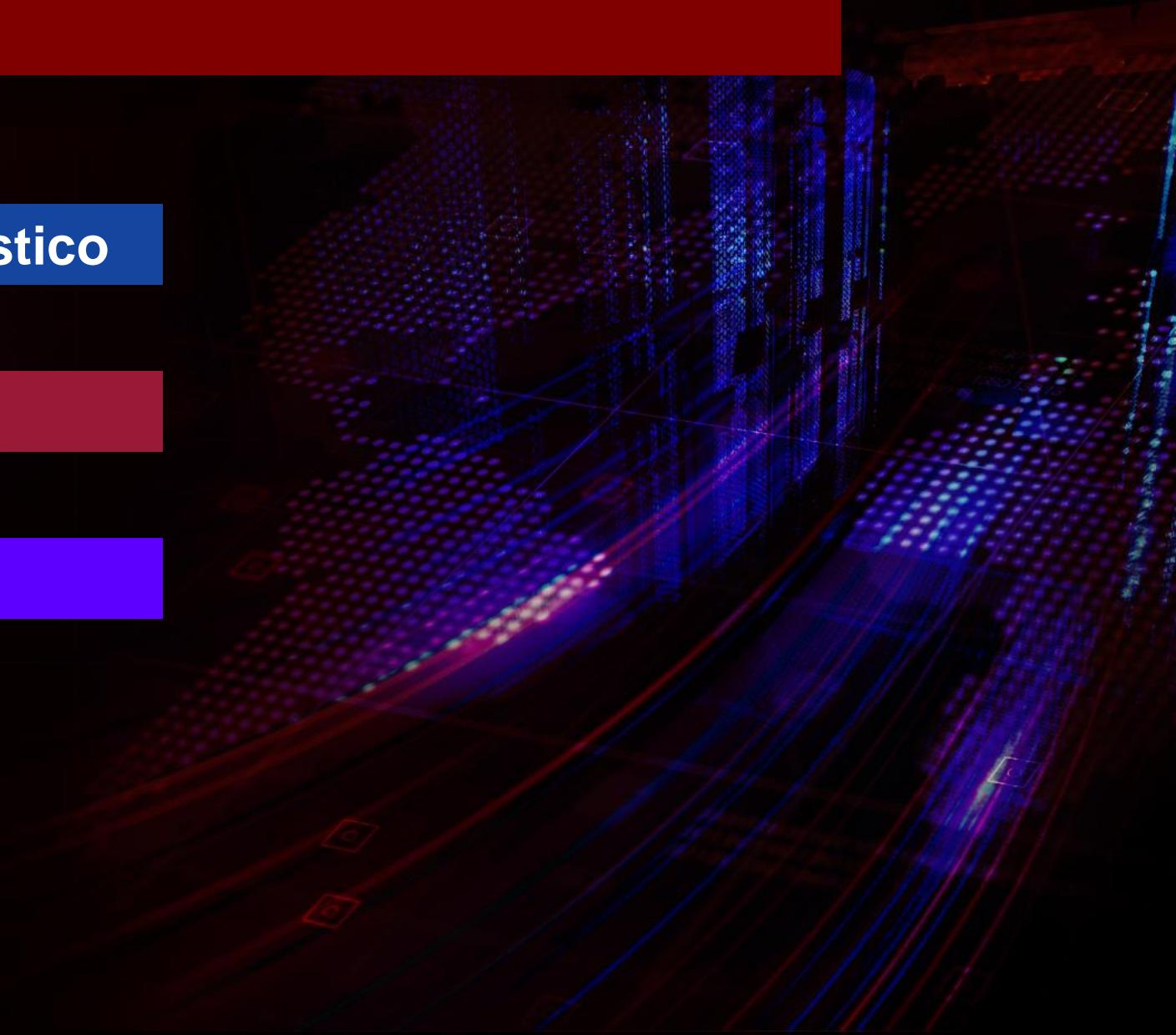


APRENDIZADO DE MÁQUINAS

Máquinas de Vetores de Suporte (SVMs)

TÓPICOS

- Teoria de Aprendizado Estatístico
- Tipos de SVMs
- Algoritmo



MARGENS



Rio Bheri, Nepal

Fonte: Madan Tiwari Mustaine, Wikimedia Commons

TEORIA DE APRENDIZADO ESTATÍSTICO (TAE)

Condições matemáticas que auxiliam na escolha de um classificador h

H : conjunto de possíveis classificadores de um algoritmo AM para um conjunto de treinamento X

Condições:

- Desempenho de h
- Complexidade

Minimizar o risco esperado (erro) de classificação

Objetivo: generalização dos resultados

TEORIA DE APRENDIZADO ESTATÍSTICO (TAE)

Conjuntos de dados grandes:
Risco empírico tende a convergir para o risco esperado

Conjuntos de dados pequenos:
Pode ocorrer a escolha de um h super ajustado

Solução:
Limitar o risco esperado

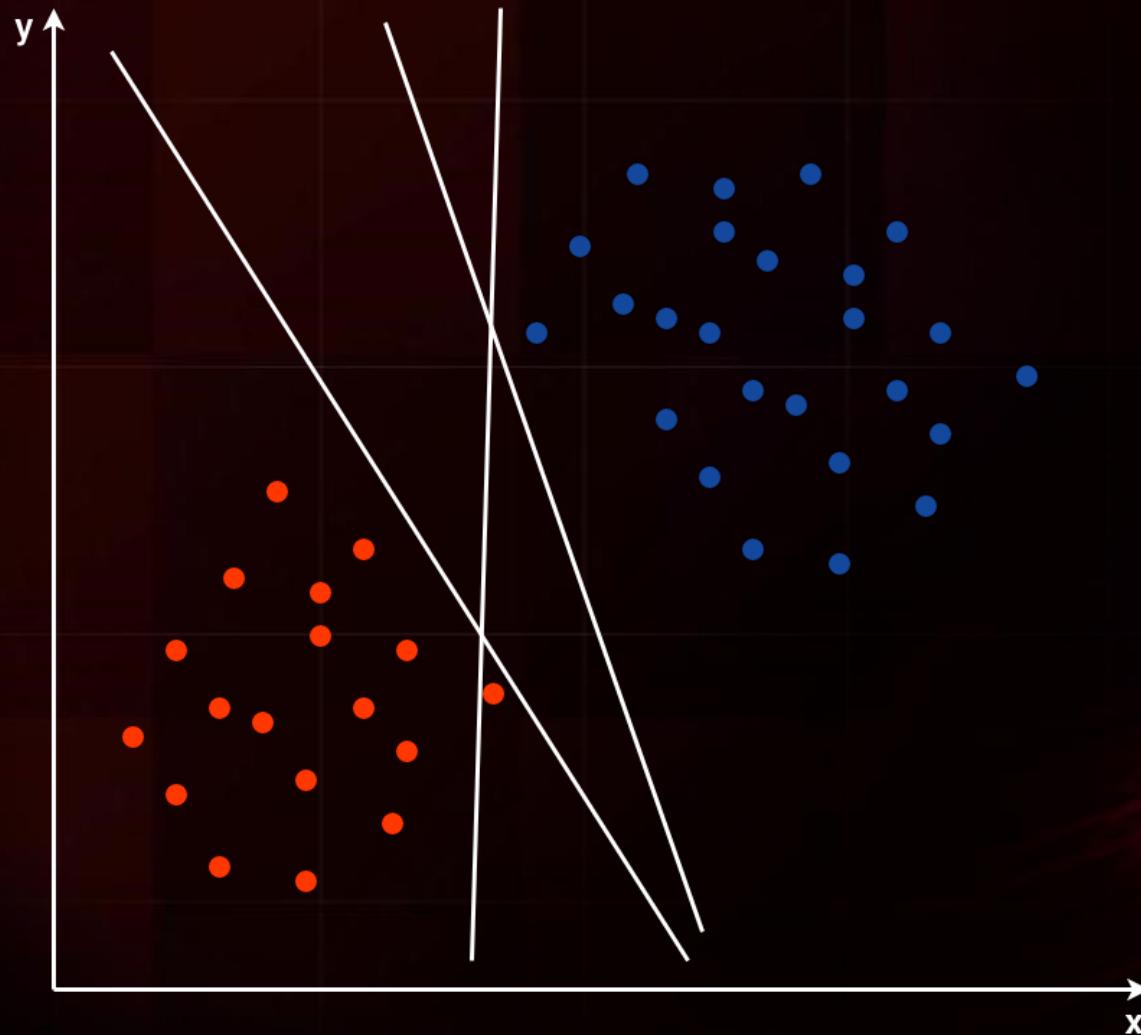
Funções de decisão lineares:
Podemos usar o conceito de margens para limitar o risco esperado



MÁQUINAS DE VETORES DE SUPORTE (SVMs)

- Aplicação da Teoria de Aprendizado Estatístico
- Classificador: separador de margem máxima
 - Fronteira linear
- Separação linear em hiperplano
 - Espaço de dimensão superior
- Opção quando não há conhecimento prévio do domínio
- Muito usadas para classificação binária
 - Podem ser usadas em domínios com mais classes

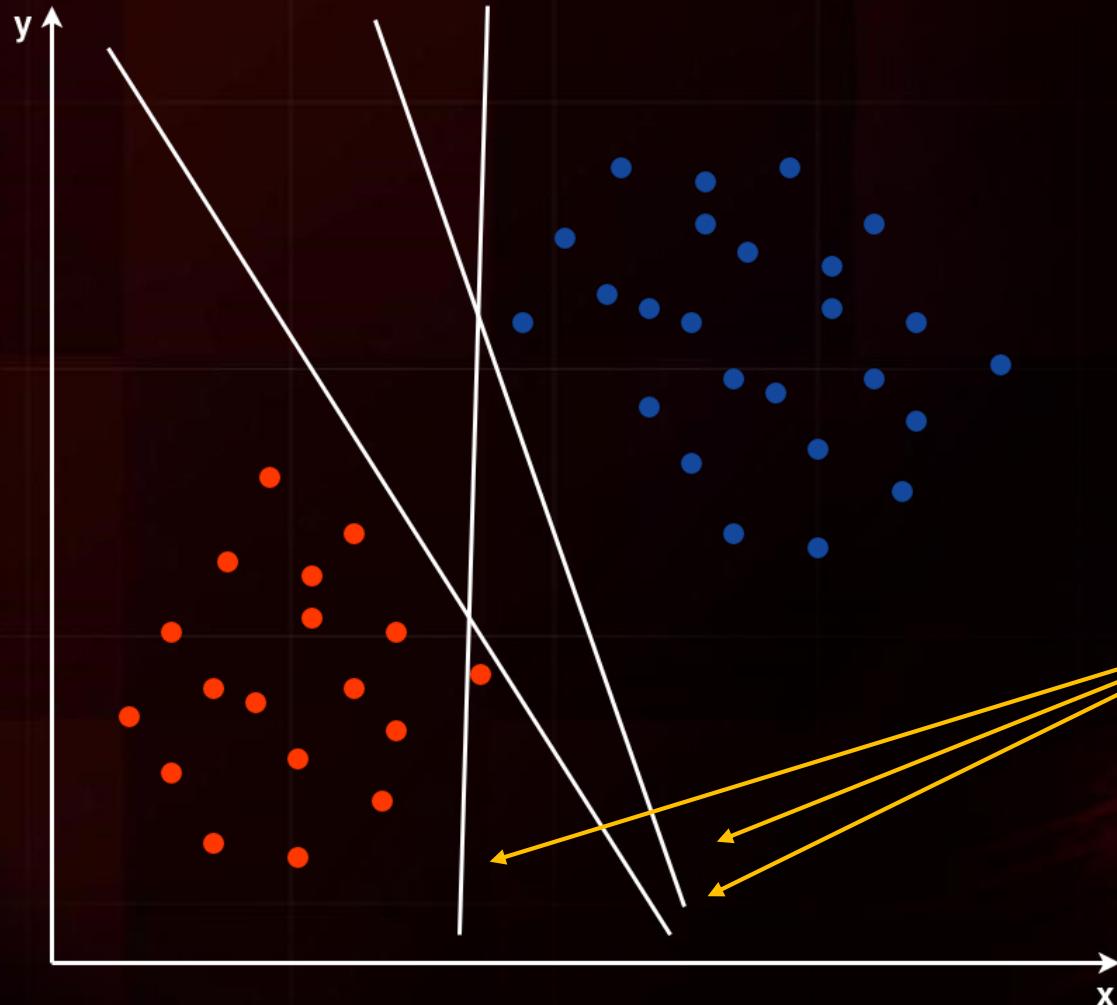
SVM: EXEMPLO



Classificação binária



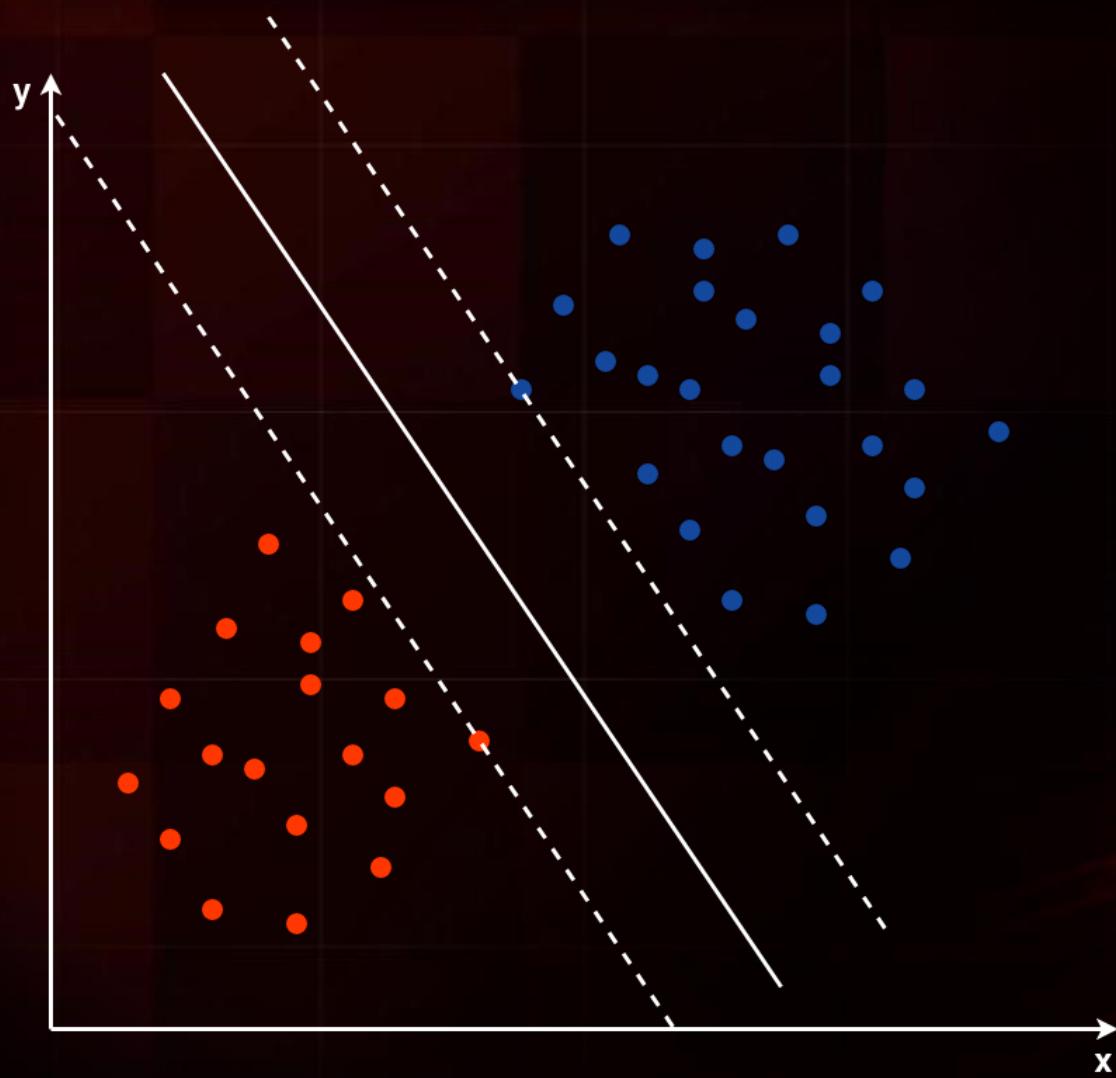
SVM: EXEMPLO



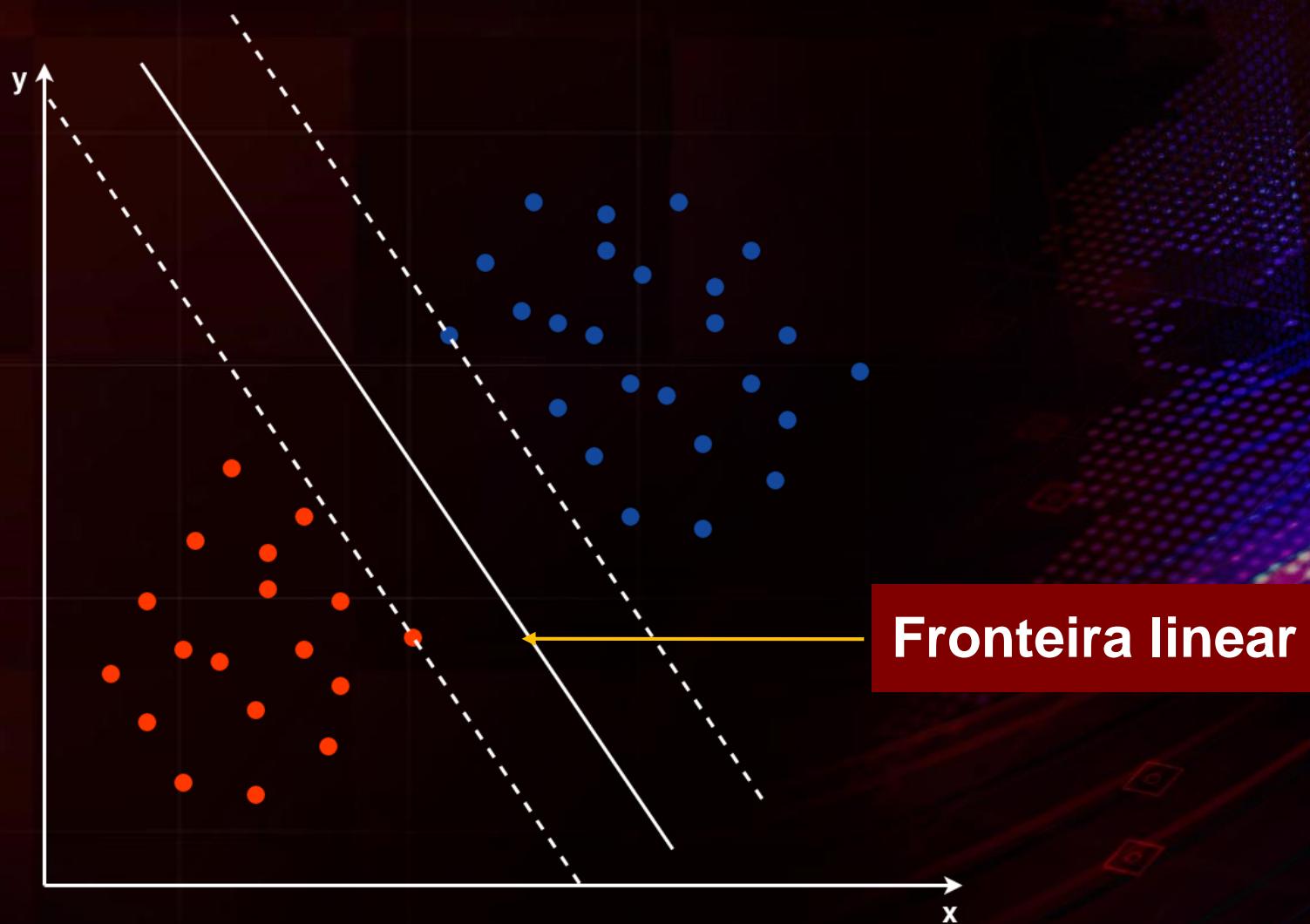
Classificação binária

Hiperplanos

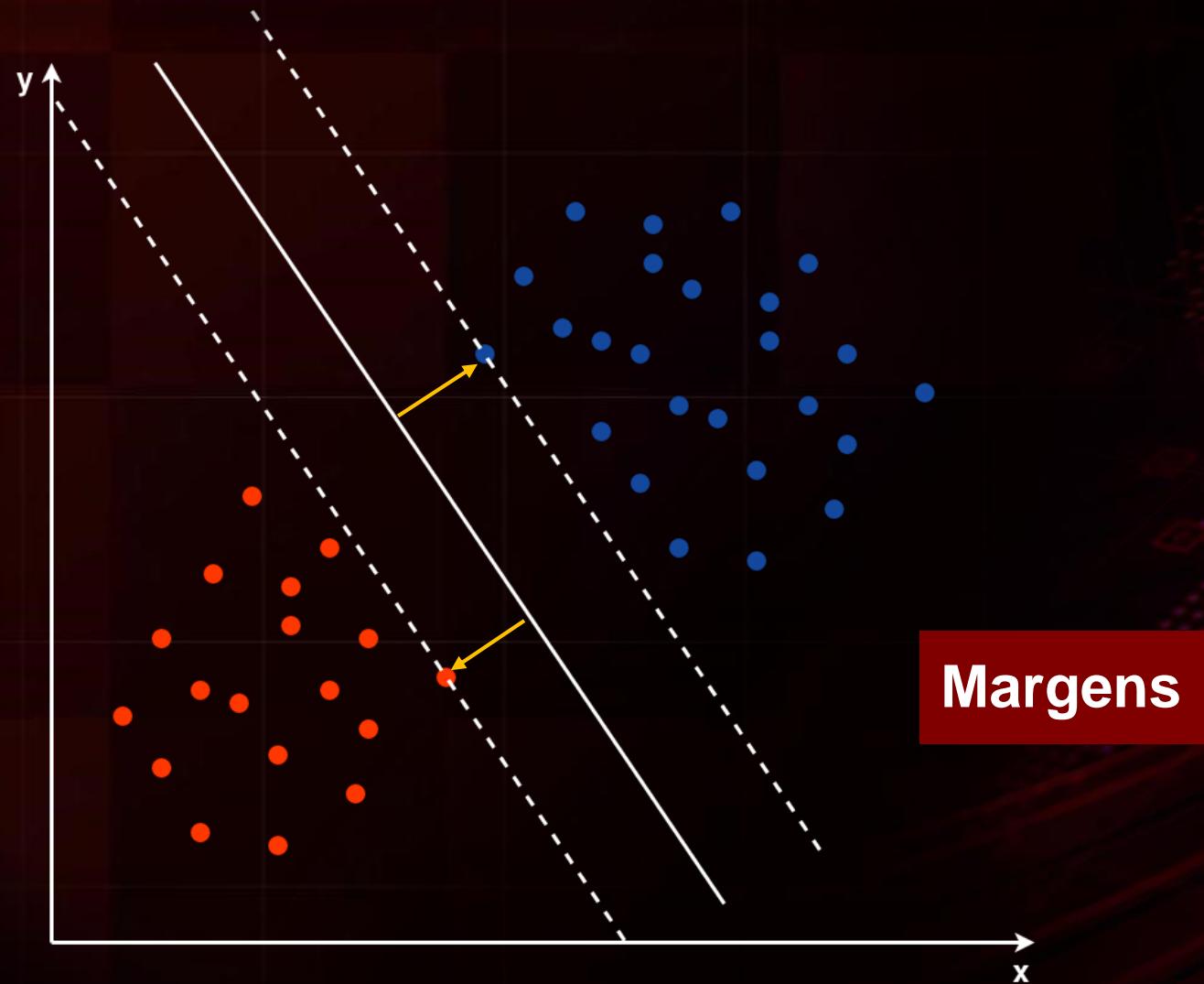
SVM: EXEMPLO



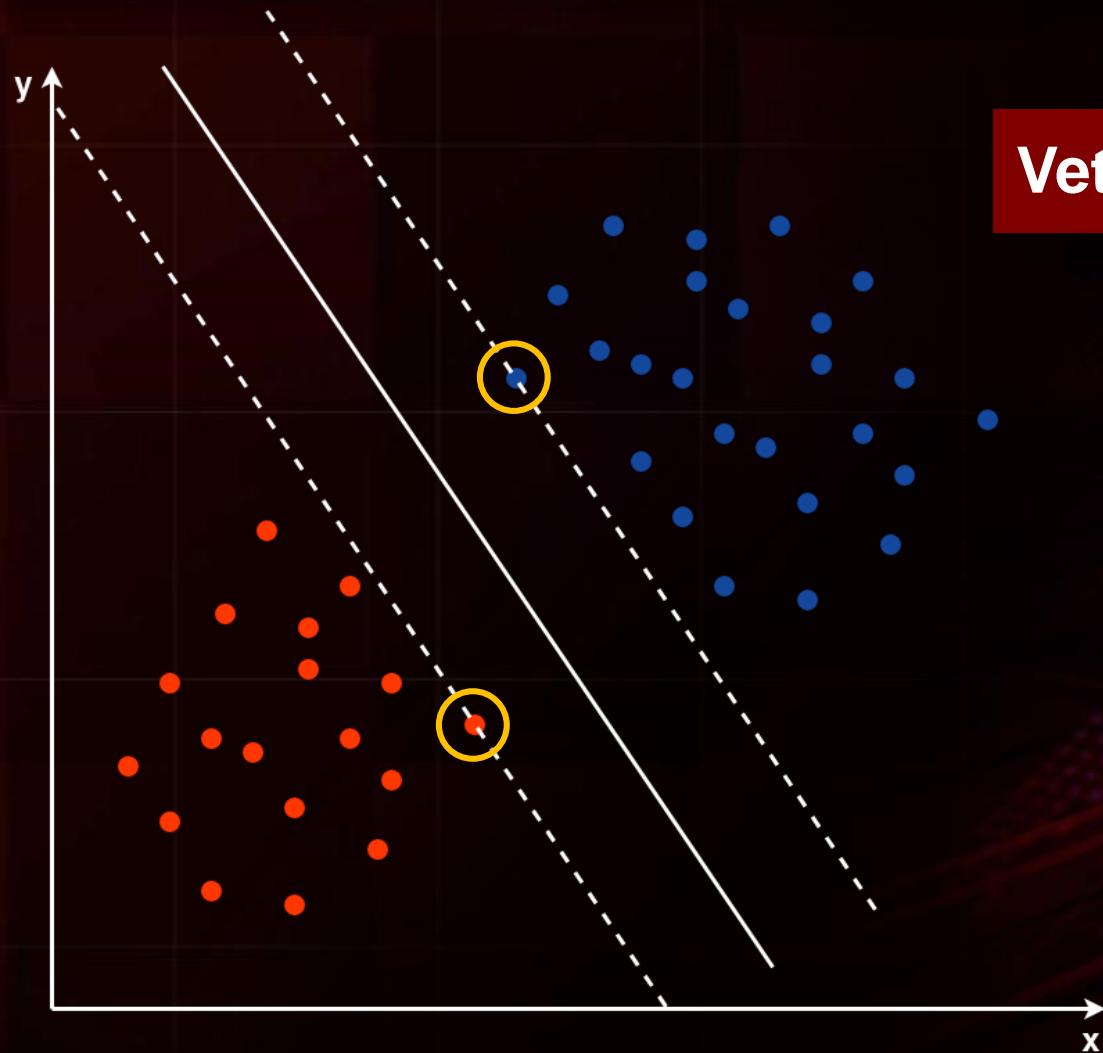
SVM: EXEMPLO



SVM: EXEMPLO

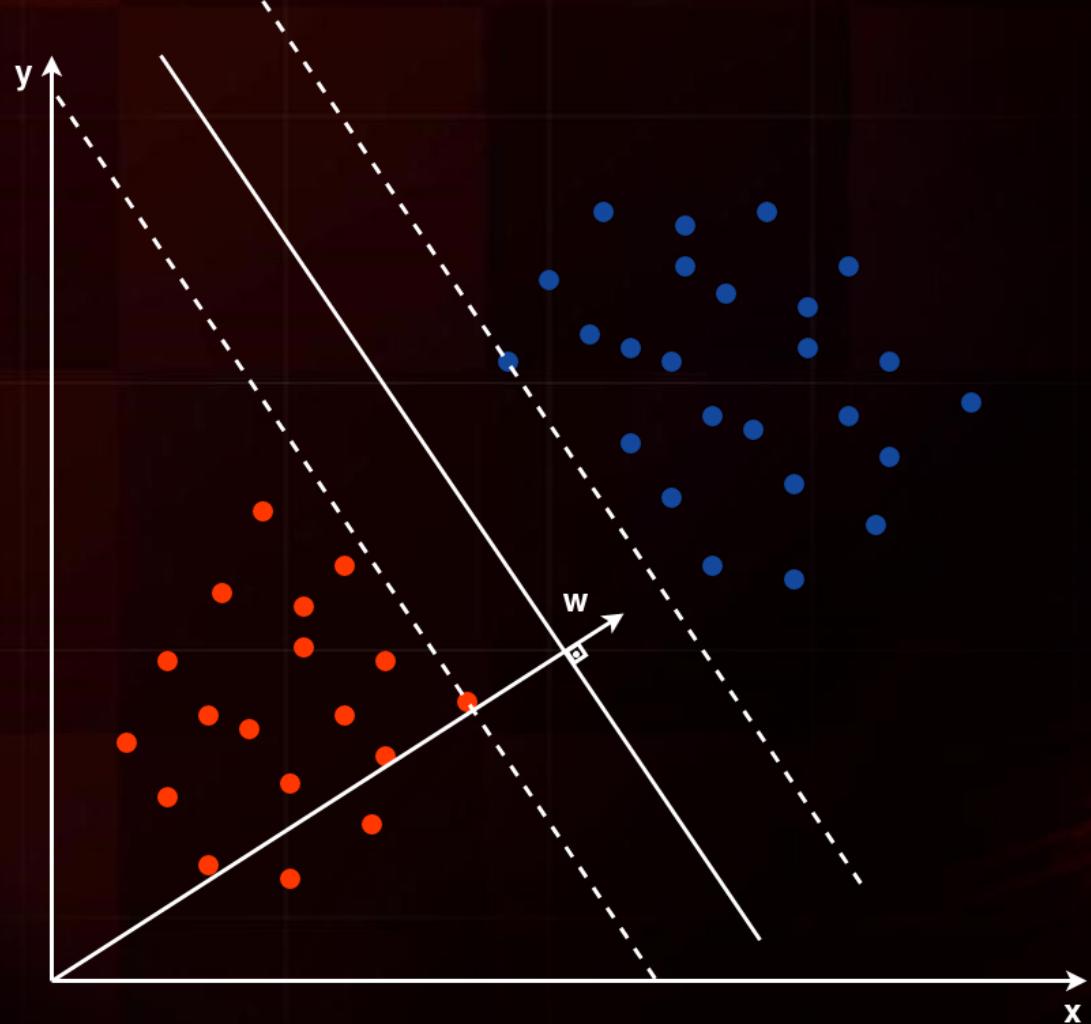


SVM: EXEMPLO



Vetores de suporte

SVM: EXEMPLO



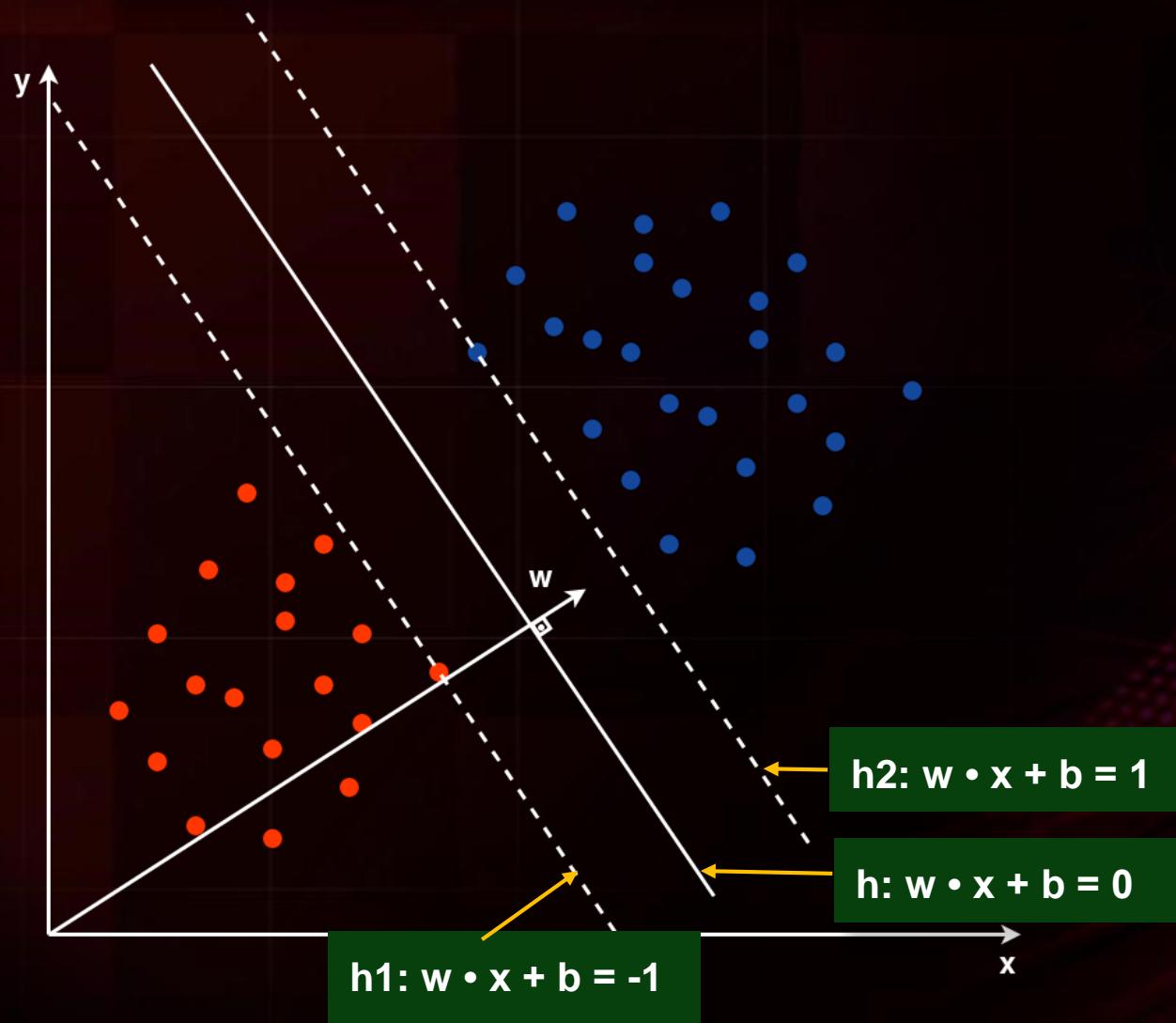
$$h(x) = w \cdot x + b$$

w: vetor normal

x: vetor de um ponto qualquer

b: constante de deslocamento de h

SVM: EXEMPLO

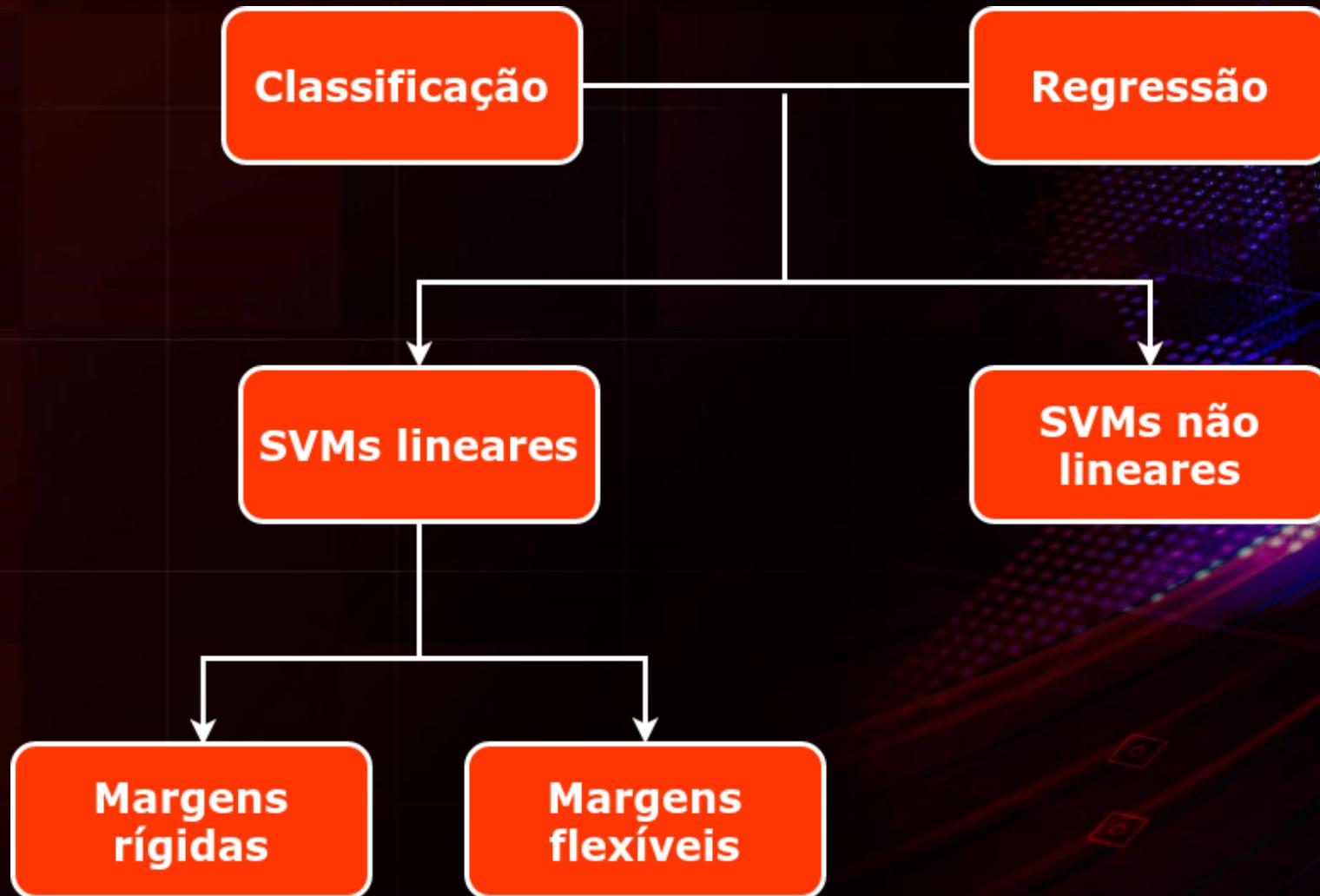


$$w \cdot x + b \geq 1$$

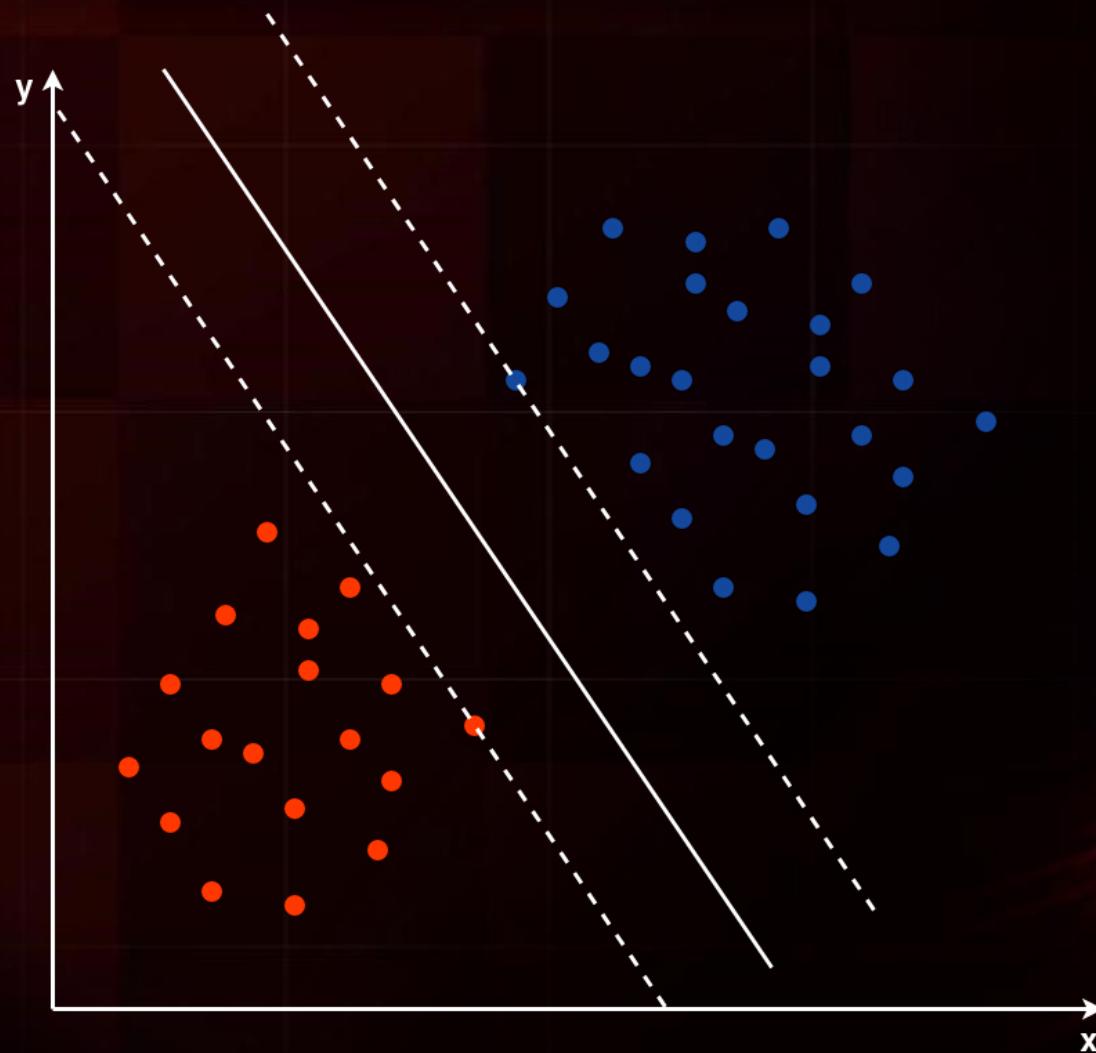


$$w \cdot x + b \leq -1$$

TIPOS DE SVMs

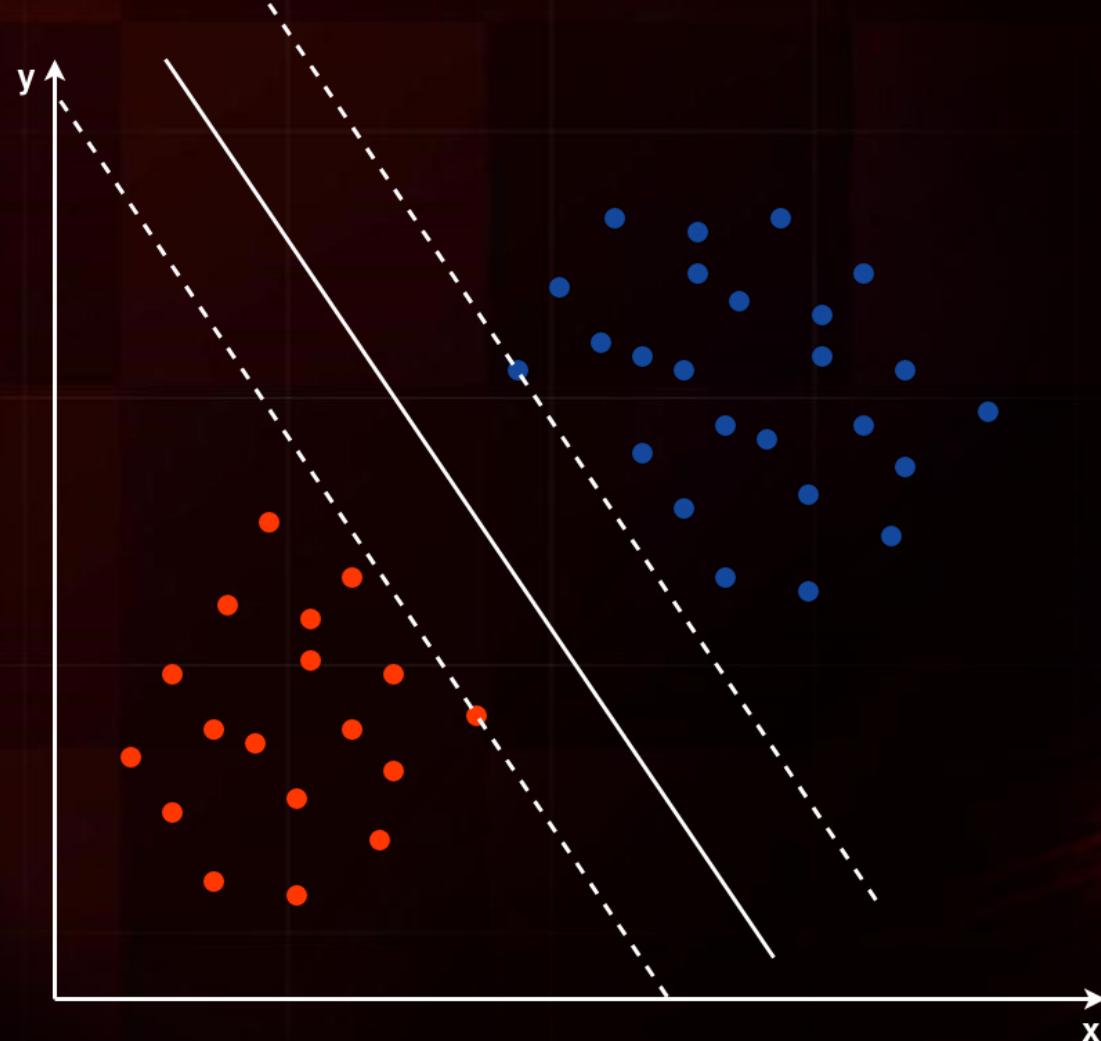


SVMs LINEARES



Problemas linearmente
separáveis

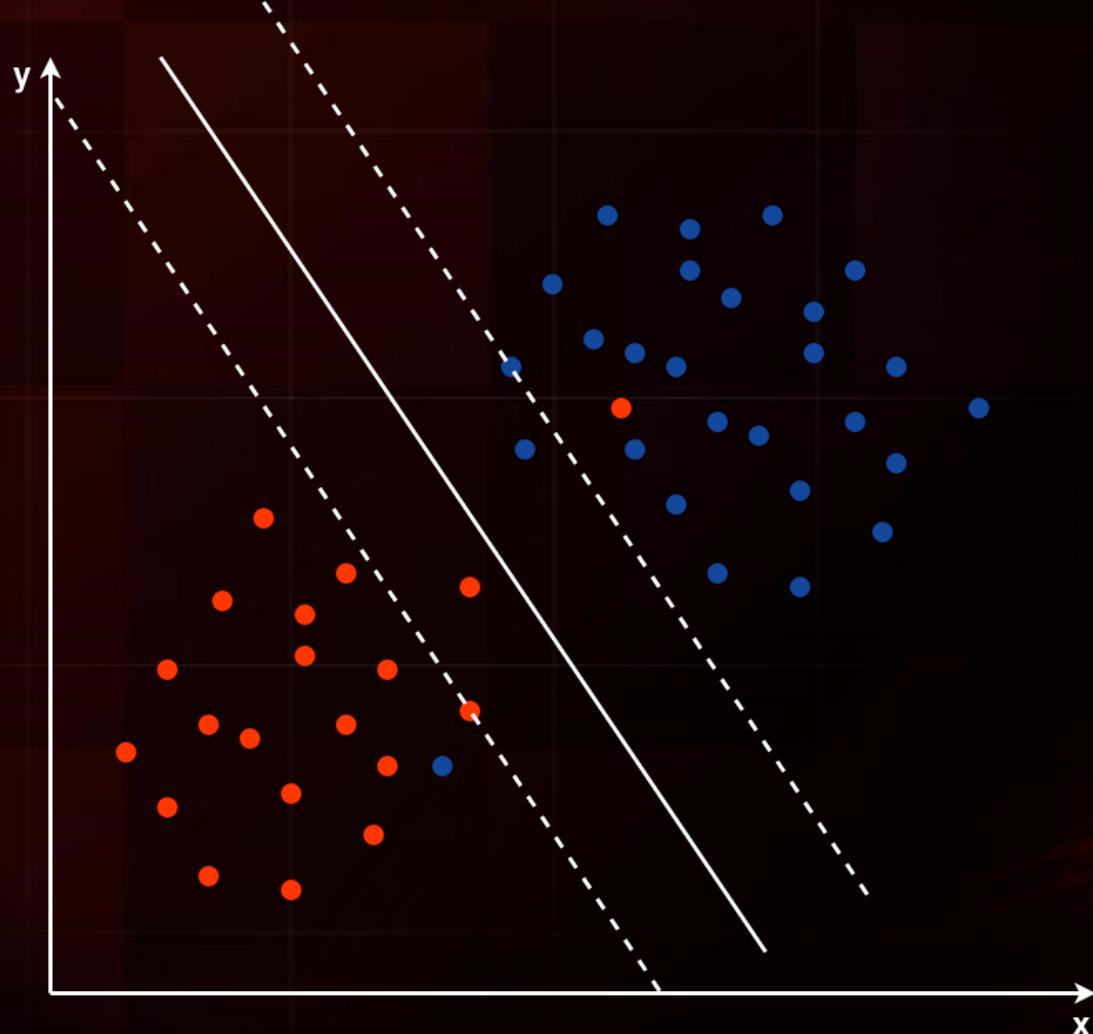
SVMs COM MARGENS RÍGIDAS



Classes: +1 ou -1

Assume que os dados podem ser totalmente separáveis

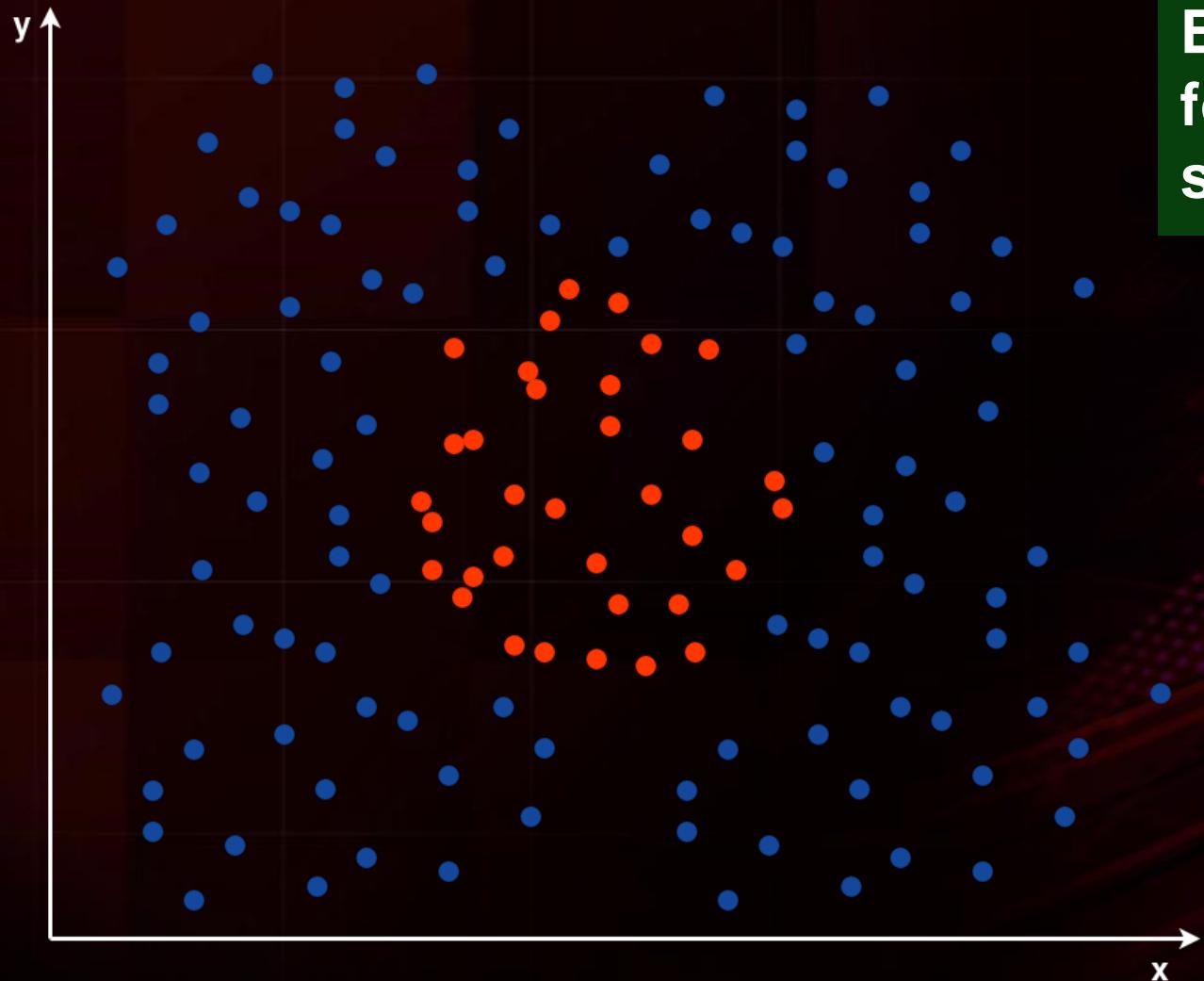
SVMs COM MARGENS SUAVES



Aceita uma taxa de erros

Mais realista: aceita outliers

SVMs NÃO LINEARES



E se os exemplos não forem linearmente separáveis?

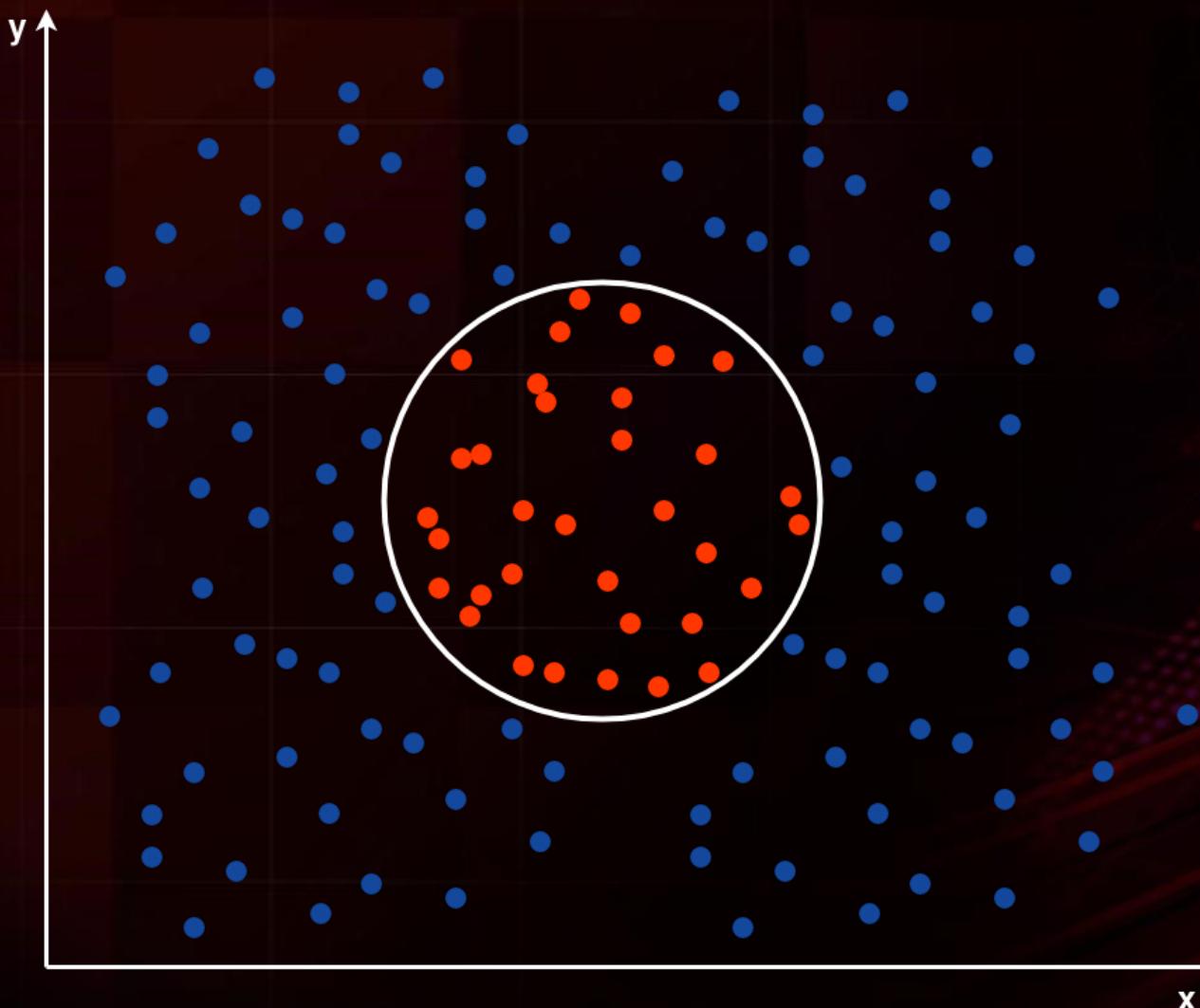
TRUQUE DE KERNEL

- **Função Kernel:**
 - **Aplicada a pares de dados de entrada**
 - **Determina produtos escalares no espaço**
 - $K(x_1, x_2) = (x_1 \cdot x_2)^2 = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$

O truque de Kernel é usado para encontrar separações lineares em espaços de dimensões variadas

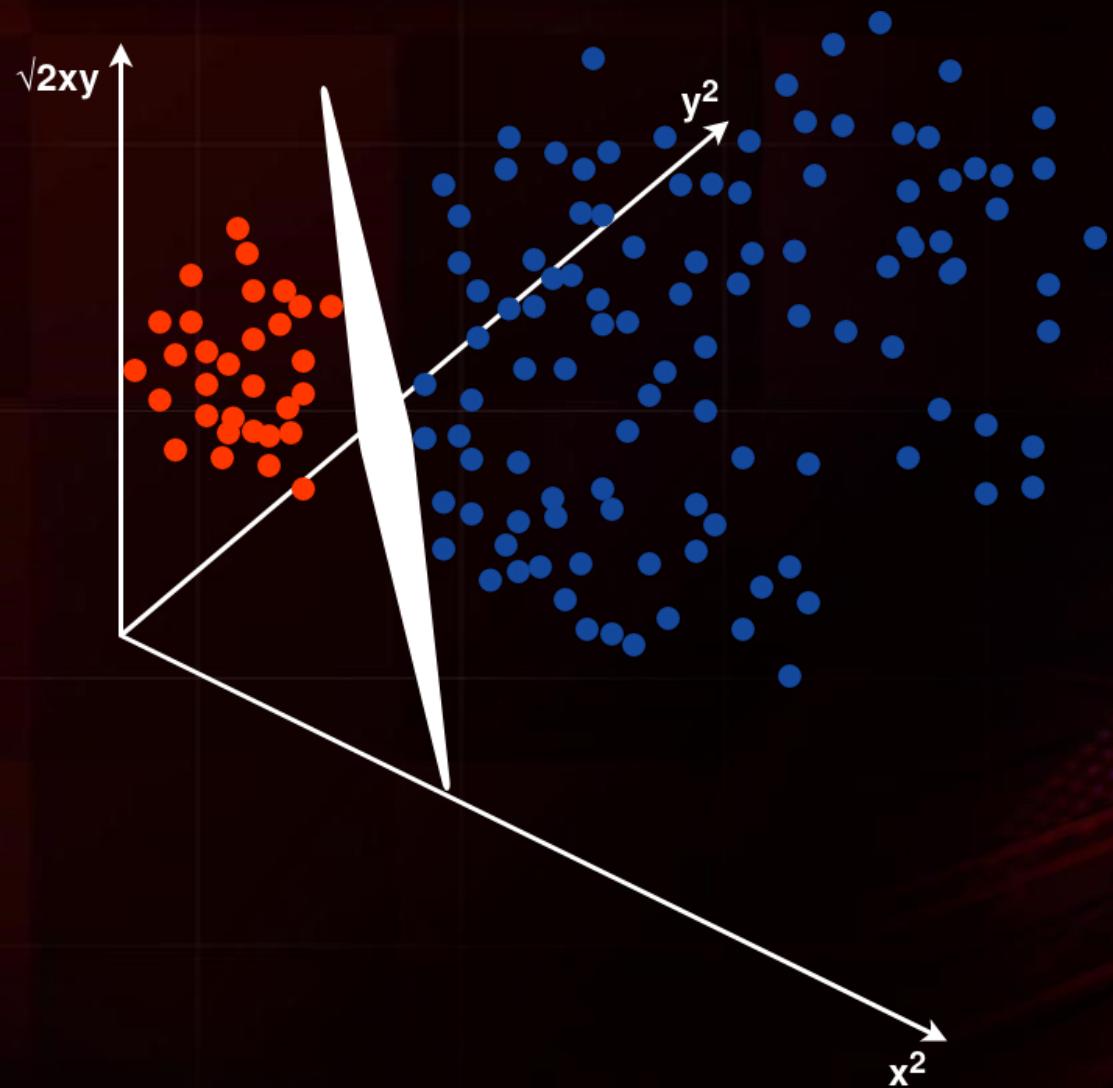


SVMs NÃO LINEARES



Transformação de
um espaço 2D...

SVMs NÃO LINEARES

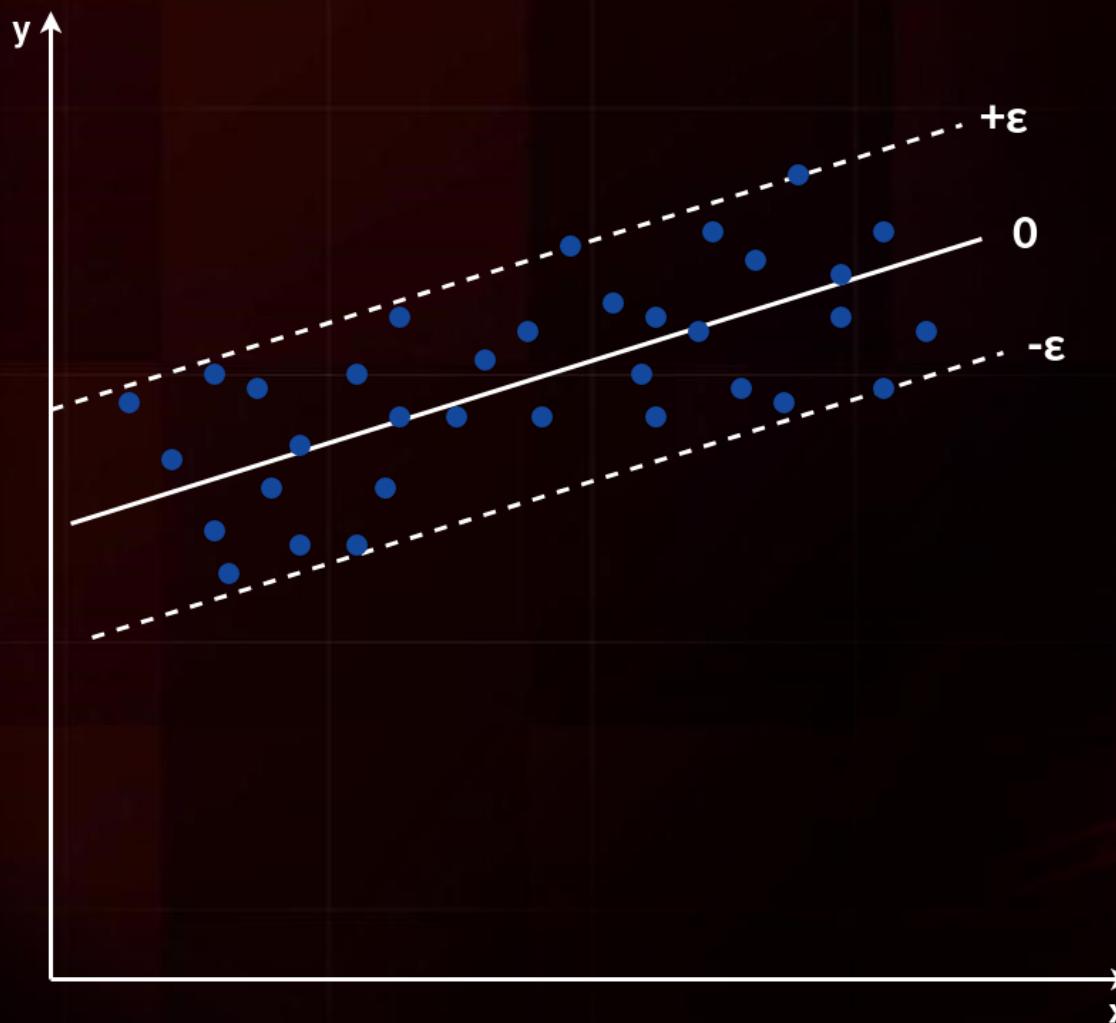


...para um espaço 3D

FUNÇÕES KERNEL

Tipo	Função $K(x_i, x_j)$	Uso
Polinomial	$(\delta(x_i \cdot x_j) + \kappa)^d$	Dados normalizados
RBF (Função Base Radial)	$\exp(-\sigma \ x_i - x_j\ ^2)$	Dados desconhecidos
Sigmodal	$\tanh(\delta(x_i \cdot x_j) + \kappa)$	Classificação binária

SVMs PARA REGRESSÃO



Algoritmo ϵ -SVM

Encontrar uma função linear
 $h(x)$ cujos dados de
treinamento desviam no
máximo de ϵ ($+ε, -ε$)

ALGORITMO*

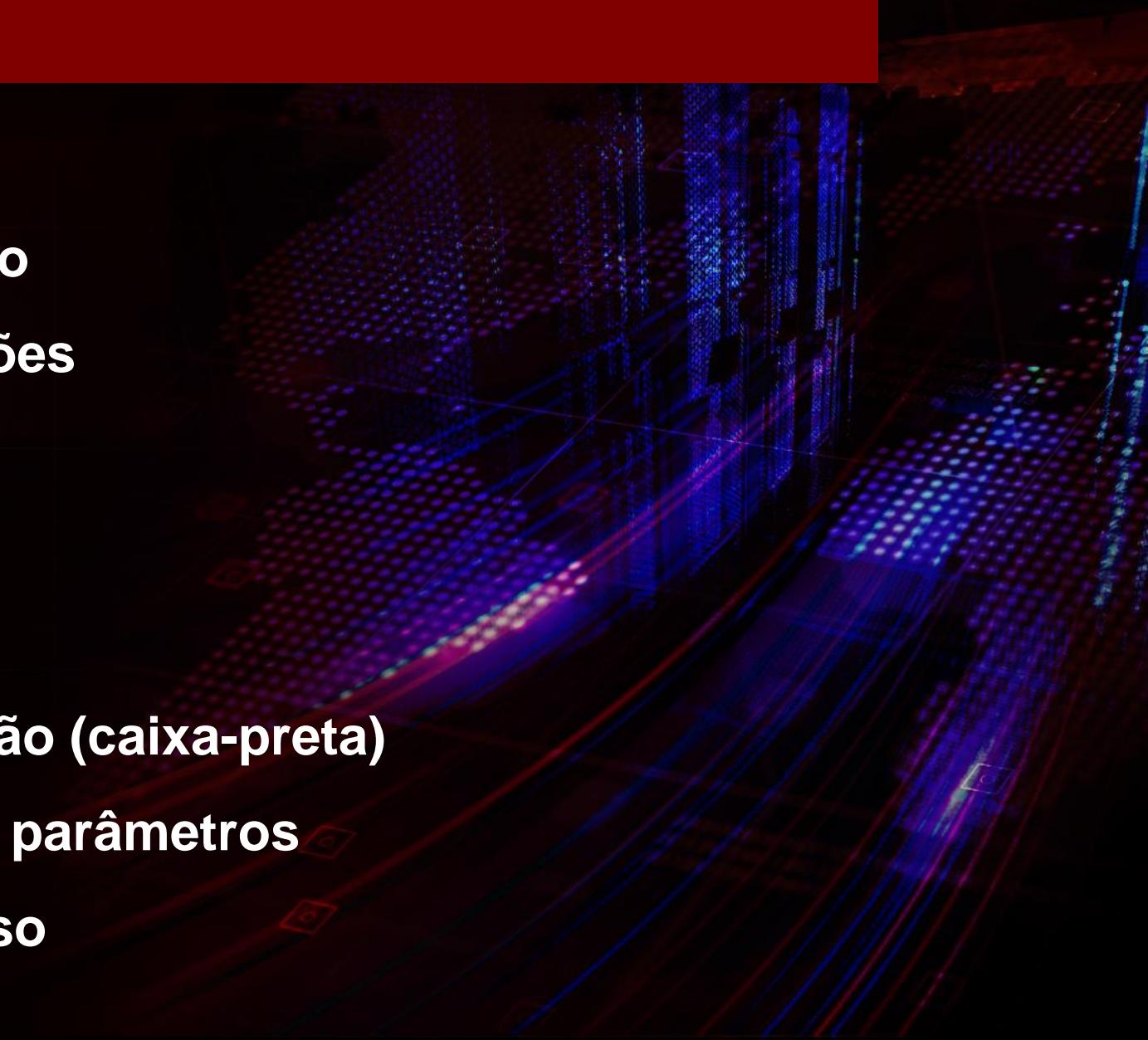
Entrada: Conjunto de dados de treinamento T
Saída: Modelo SVM treinado

- 1. Para cada solução possível (classificador h)**
- 2. Encontre os vetores de suporte (T)**
- 3. Calcule a margem do classificador h**
- 4. Escolha h que maximiza a margem entre as classes**

* SVM linear

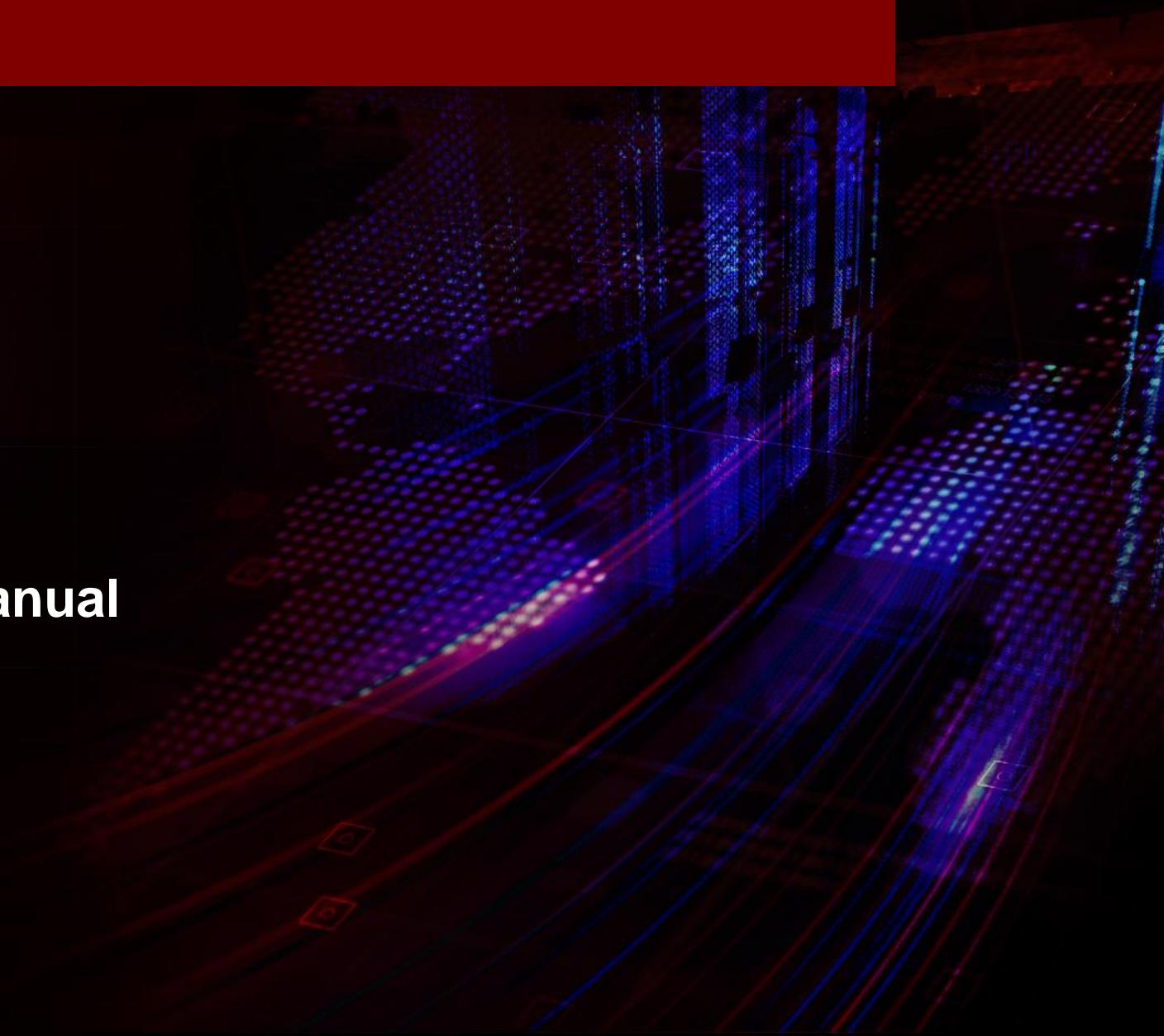
CONSIDERAÇÕES

- **VANTAGENS:**
 - Capacidade de generalização
 - Objetos de grandes dimensões
 - Determinístico
 - Robusto a ruídos
- **DESVANTAGENS:**
 - Modelo de difícil interpretação (caixa-preta)
 - Sensibilidade à escolha dos parâmetros
 - Computacionalmente custoso



APLICAÇÕES

- Classificação de imagens
- Classificação textual
- Reconhecimento facial
- Reconhecimento de escrita manual
- Classificação de genes
- Análise de séries temporais



REFERÊNCIAS

- Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina. **Katti Facelli et al. SAGAH, 2021.**
- Inteligência Artificial. **Peter Norvig. Grupo GEN, 2013.**