

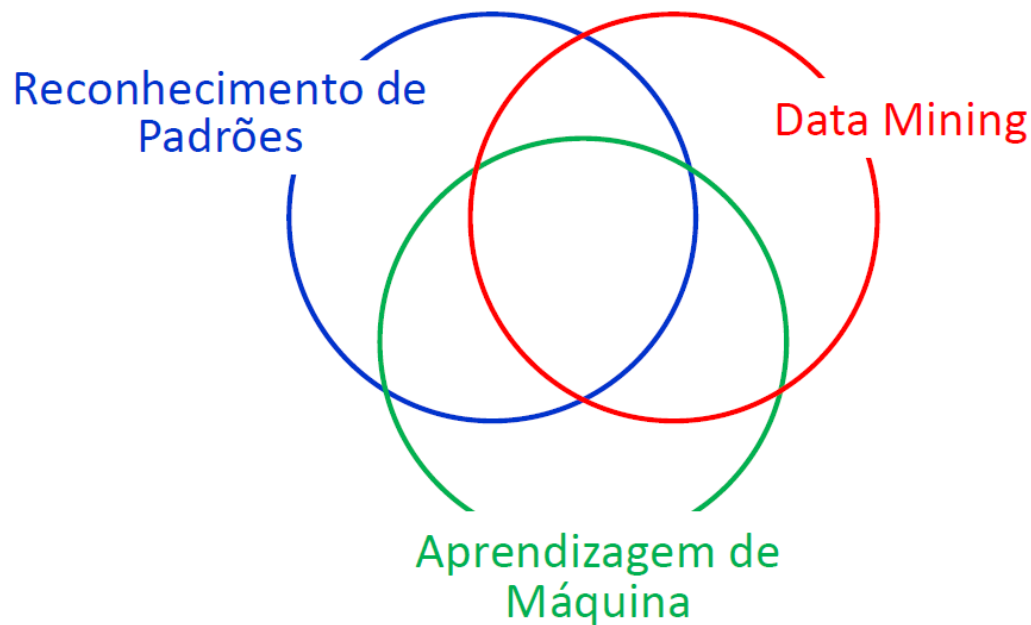
Machine Learning

Prof. Angelo Orsatto

Introdução

Parte da Inteligência Artificial

Áreas conectadas e
Interligadas



Reconhecimento de Padrões

Tem como uma de suas principais aplicações atribuir a um determinado objeto, uma classe entre várias possíveis

Este processo de rotulação recebe a alcunha de classificação

O procedimento consiste em analisar um conjunto de informações (vetor de características) sobre o elemento a ser classificado e, segundo critérios definidos, determinar a qual classe ele pertence

Os responsáveis por realizar a atribuição de rótulos aos elementos ainda não classificados são os classificadores

Reconhecimento de Padrões

Padrão (instância):

- é uma entidade, objeto, processo ou evento, vagamente definido, que pode assumir um nome

Classe:

- conjunto de padrões que possuem características em comum

Característica ou Atributo:

- dado extraído de uma amostra por meio de medida e/ou processamento. Em geral são organizadas na forma de um vetor de características

Reconhecimento de Padrões

Classificação:

- atribuir classes para as amostras, baseado em suas características

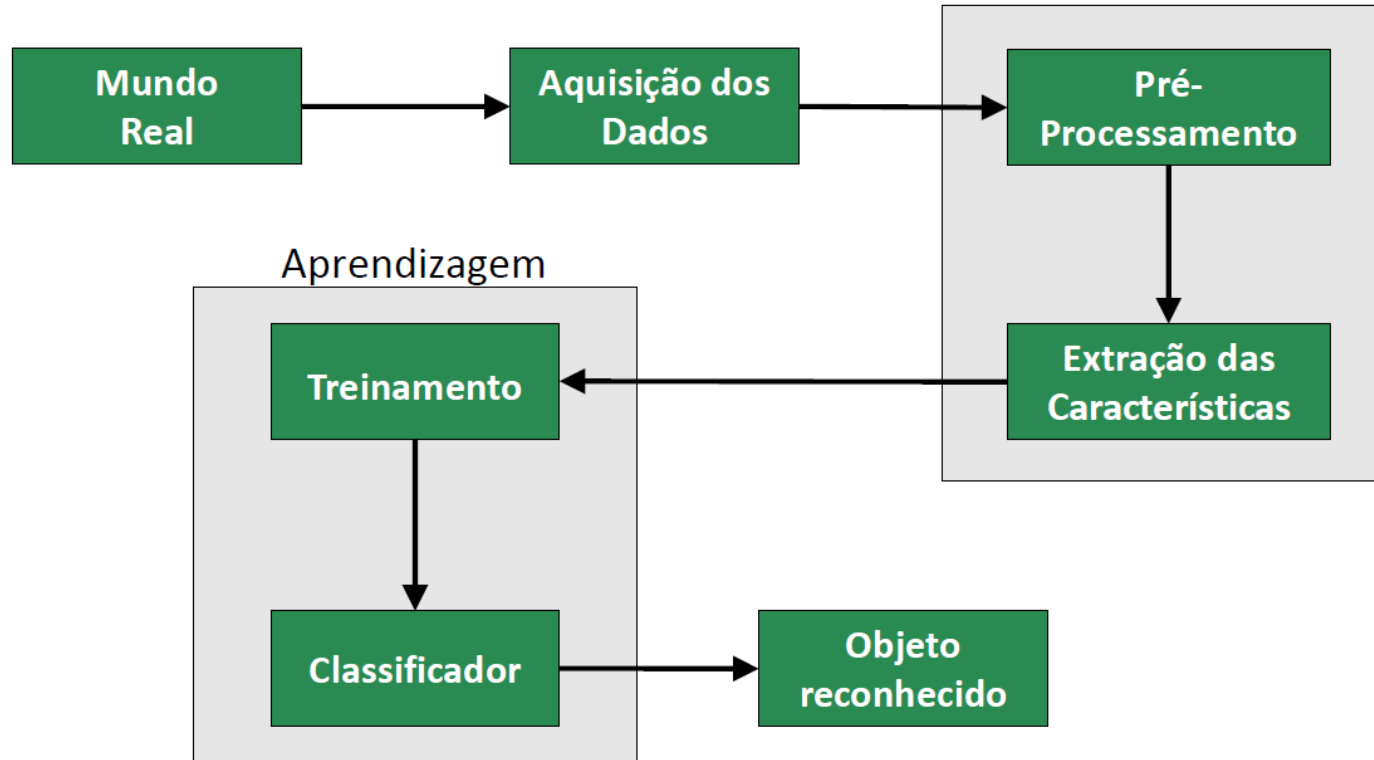
Classificadores:

- utilizados para classificar ou descrever padrões ou objetos a partir de um conjunto de propriedades ou características

Ruído:

- distorção, falha ou imprecisão que ocorre na aquisição dos dados

Sistema de Reconhecimento



Sistema de Reconhecimento

Mundo real - Aquisição dos dados

Varia de acordo com o problema

Totalmente dependente da aplicação e objetivo

Envolve, necessariamente o conhecimento do especialista no processo de identificar o que deve ser analisado

Sistema de Reconhecimento

Pré processamento - Extração das Características

Conhecimento do especialista para escolher as características adequadas

Remoção de ruídos

Duplicatas, Outliers

Tratamento específicos

Processamento de imagens

Treinamento

Quanto mais robusto o treinamento, melhor tende a ser o resultado

No treinamento é necessário que todas as classes estejam presentes

Manter a proporção original das classes

Aprendizagem de Máquina

Capacitar um modelo (classificador) através do treinamento, a reconhecer uma nova instância com base no conhecimento prévio do problema

Espera-se que a aprendizagem seja a mais efetiva e generalista possível, alcançando a maior taxa de reconhecimento

Exemplo

Reconhecimento automático de peixes:

Salmão



Tilápia



Exemplo

Aquisição

- Câmera para capturar imagens dos peixes
- Uso de ambiente controlado

Pré Processamento

- Remoção do fundo
- Remoção de ruídos
- Equalização do histograma
- Segmentação de cada peixe de forma individual

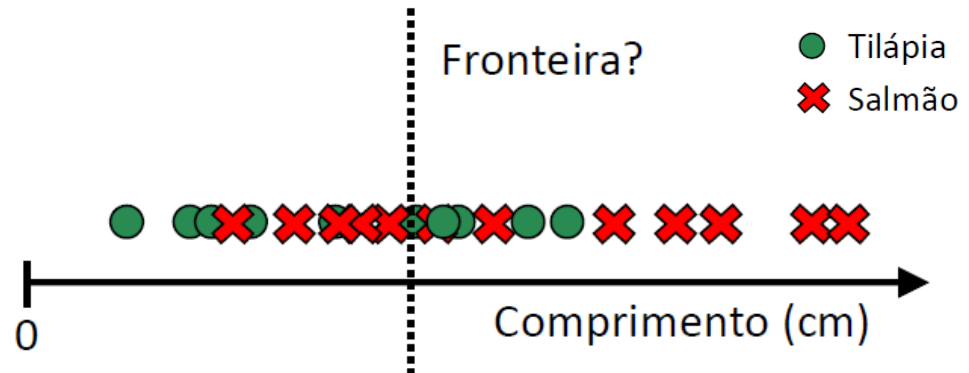
Exemplo

Extração das características

Com base na imagem segmentada, o sistema deve levantar as características definidas pelo especialista do domínio

Ex.: adotar o comprimento dos peixes

Comprimento (cm)	Classe
25	Tilápia
35	Tilápia
38	Salmão
49	Salmão
...	...



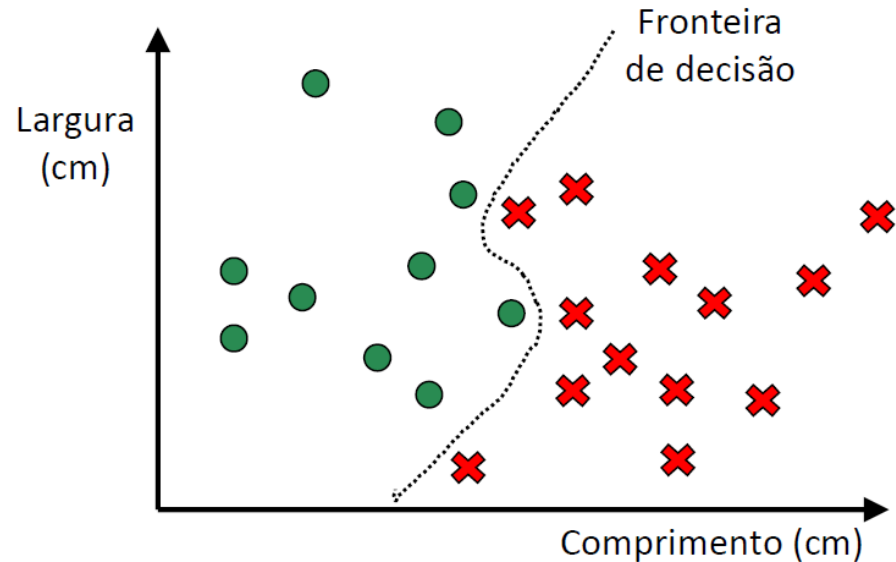
Exemplo

Extração das características

Ex.: adotar o comprimento e largura dos peixes

Comprimento (cm)	Largura (cm)	Classe
25	19	Tilápia
35	24	Tilápia
38	15	Salmão
48	18	Salmão
...

● Tilápia
✗ Salmão



Características

Tipos de características

Numéricas (float): comprimento, largura

Nominal (enumerado): tipo de escama, posição nadadeira

Inteiros (tratados como numérico): número de barbatanas, nadadeiras

O tipo da características influencia na escolha e desempenho do classificador

Características

Discriminante

- Identificar fatores que podem identificar o objeto exclusivamente

Robusta

- Capaz de absorver variações dentro da classe

Invariantes

- Presentes em todas as instâncias

Possibilidade de conversão em valores

- Atributo pode ser extraído e convertido em valor

Classificadores

KNN

Árvores de decisão

Support Vector Machine

Multilayer Perceptron

KNN

K – Nearest Neighbors

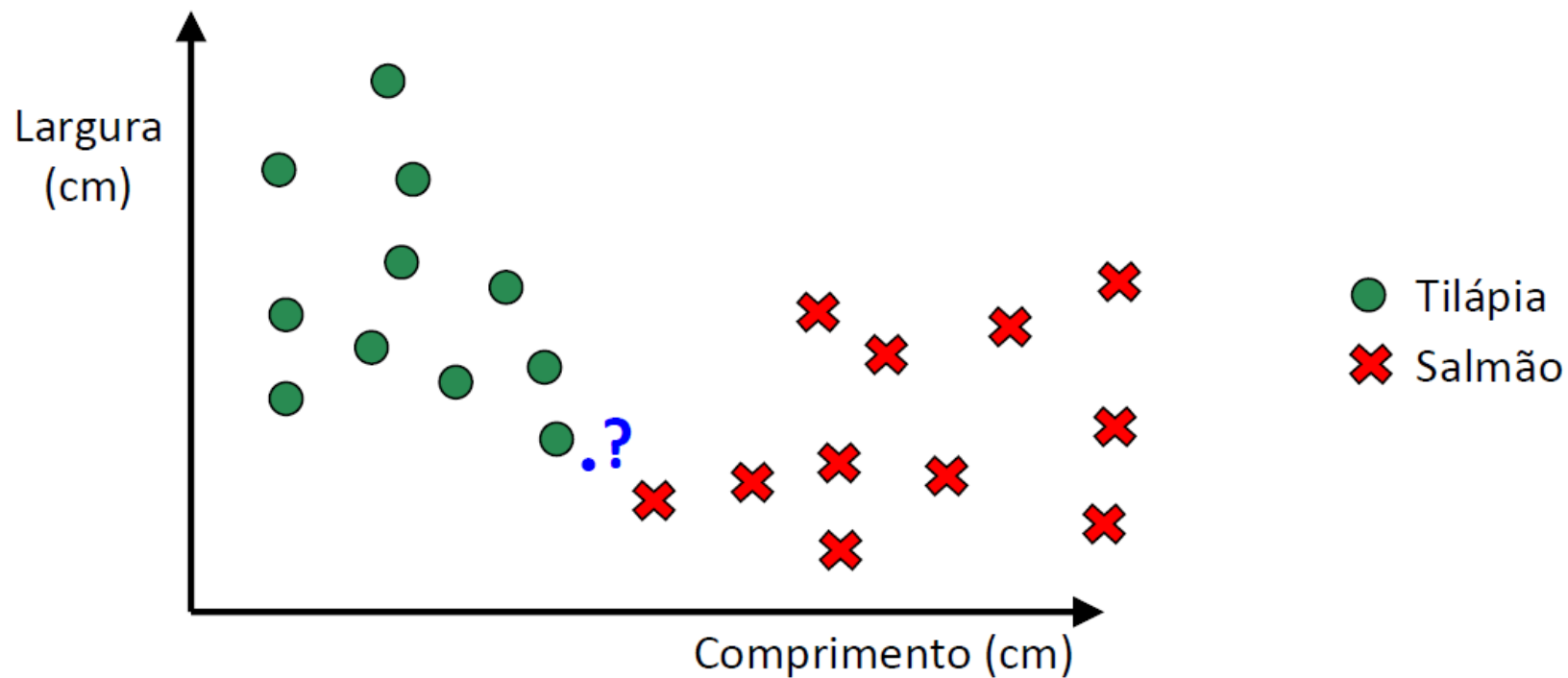
K vizinhos mais próximos

Projeta a instância de teste no espaço de características do conjunto de treino

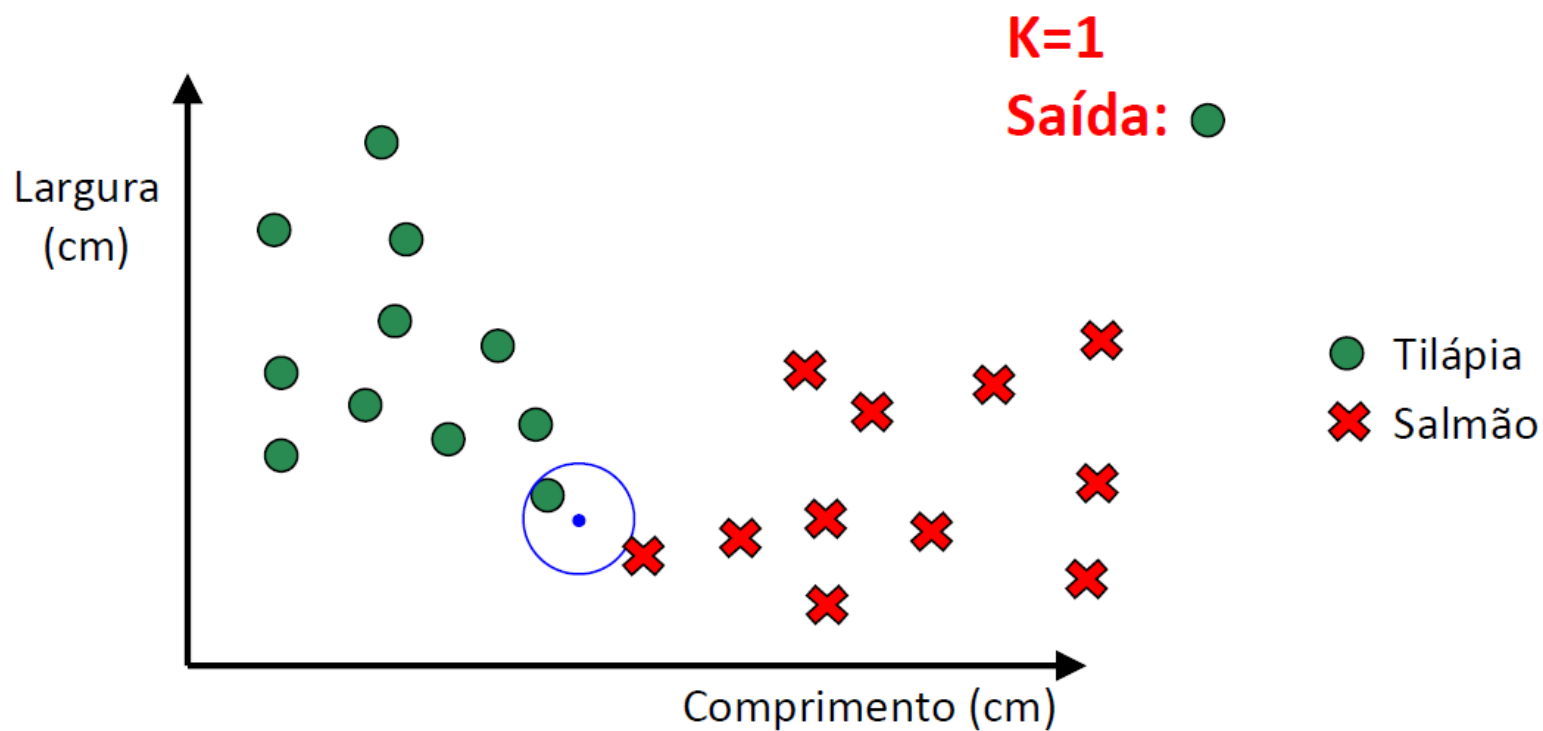
Seleciona os K elementos mais próximos dentro deste espaço

Atribui a classe que for mais presente dentro do grupo formado pelos K vizinhos mais próximos

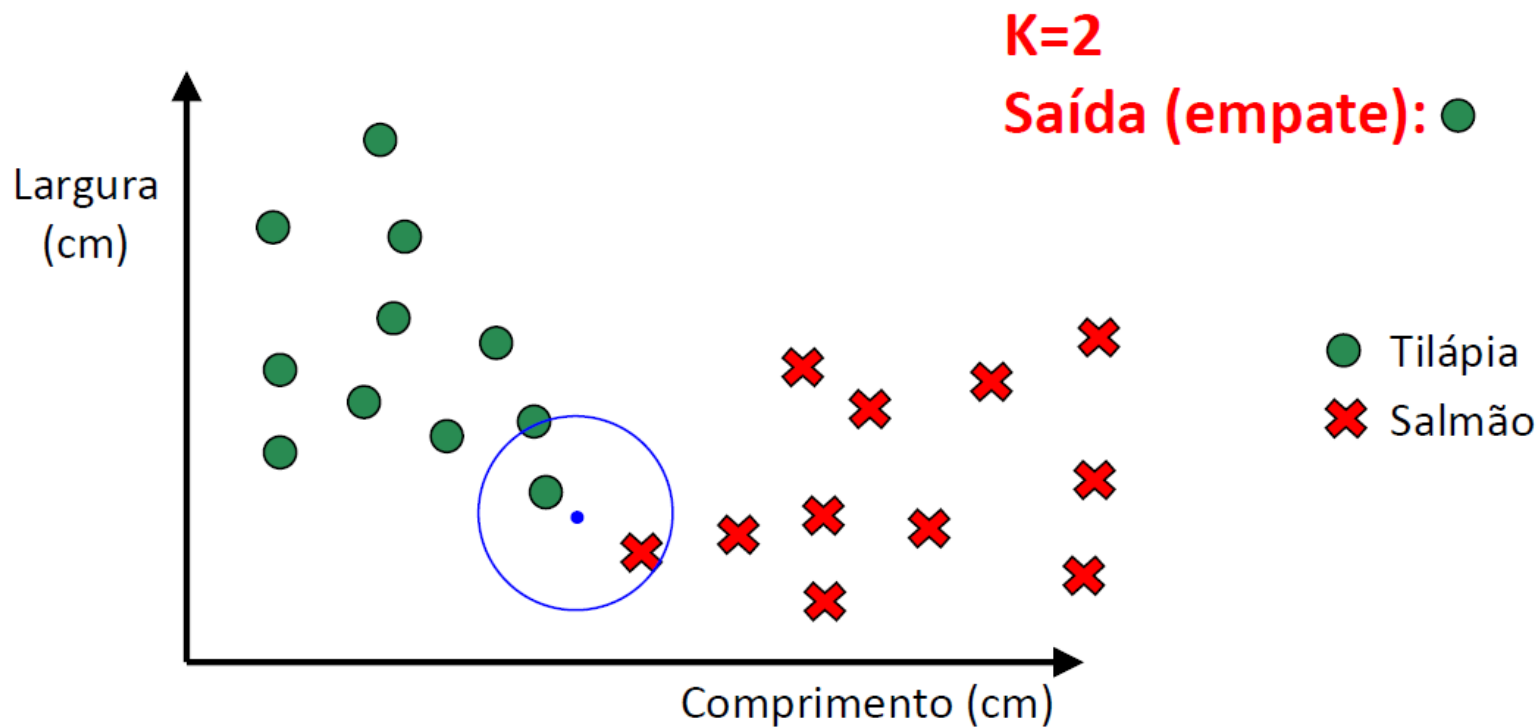
KNN



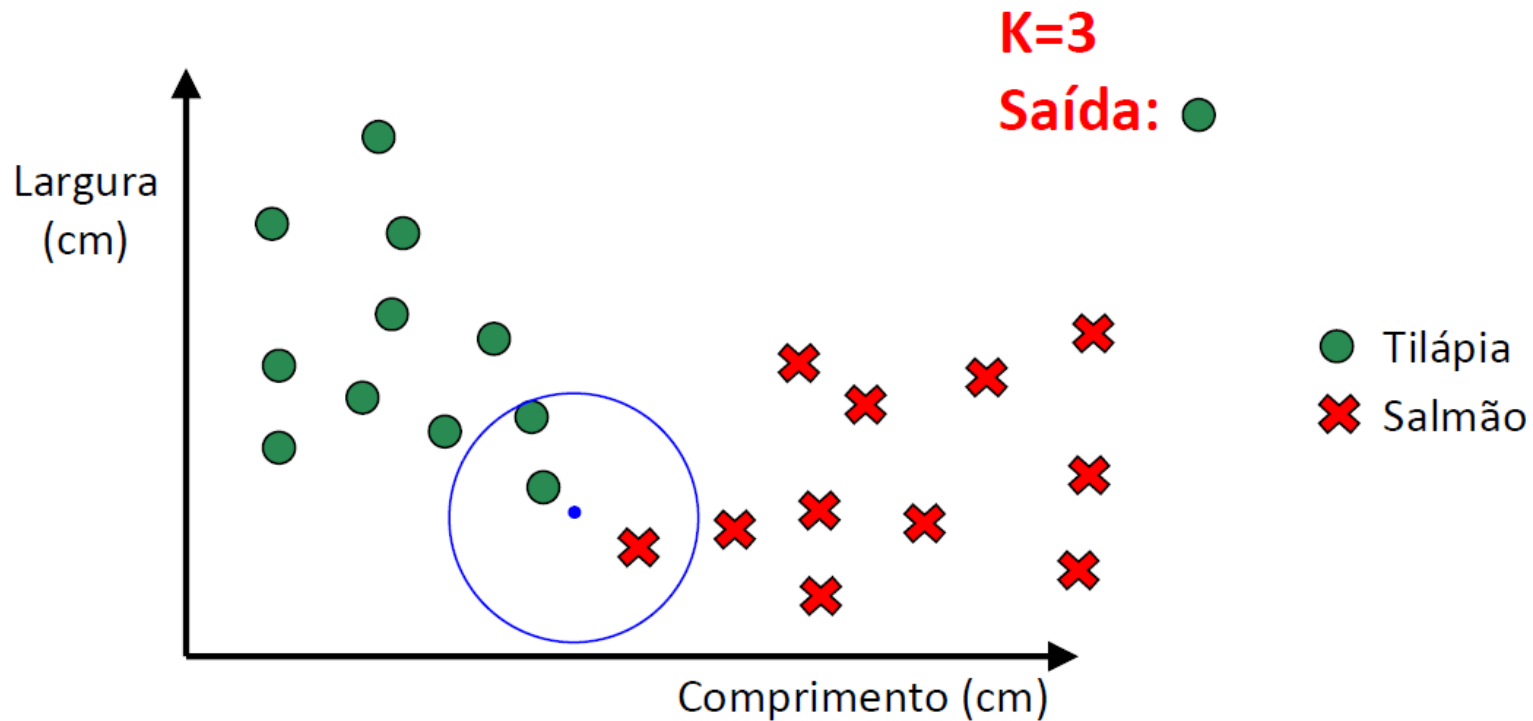
KNN



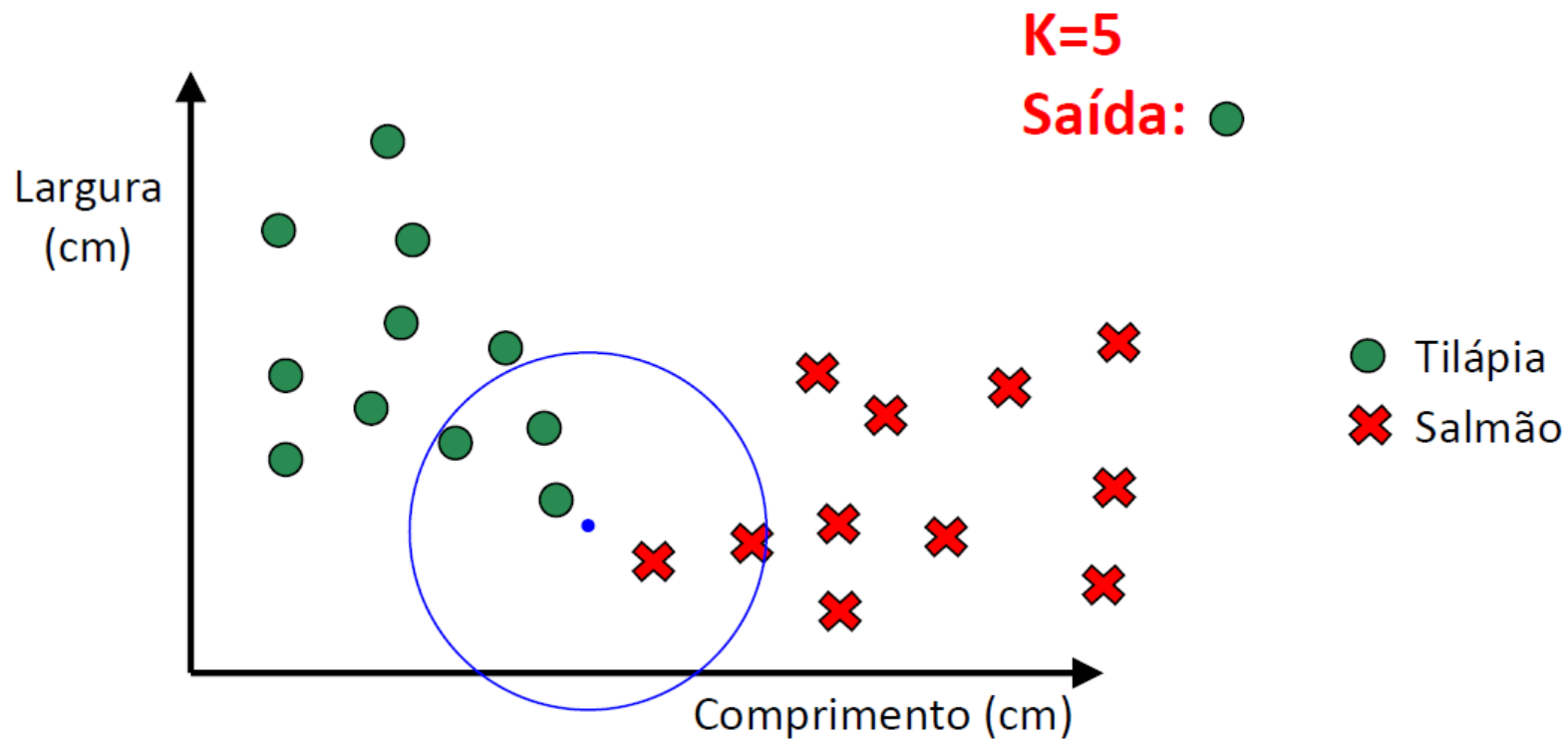
KNN



KNN



KNN



KNN

Treino rápido, etapa teste nem tanto

Todo custo do método fica praticamente na classificação

Cada instância de teste que chega deve-se procurar (calcular semelhança) com todas as instâncias do treino

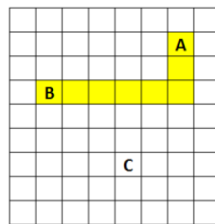
KNN

Métricas de proximidade:

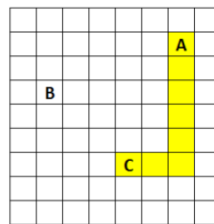
Inverso da distância Euclidiana

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

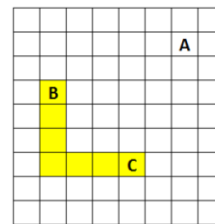
Distância de Manhattan



Distância = 6



Distância = 6



Distância = 5

-Quando menor a distância, mais próximo do teste

KNN

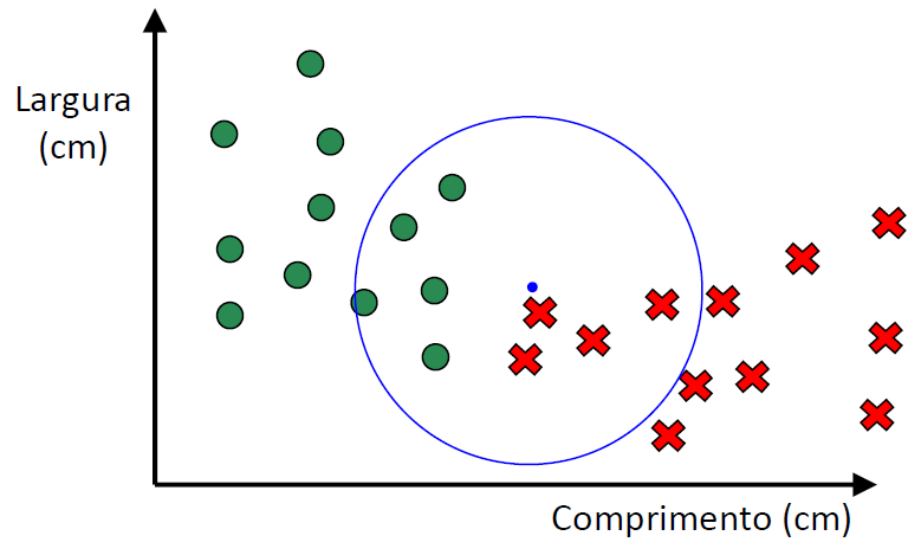
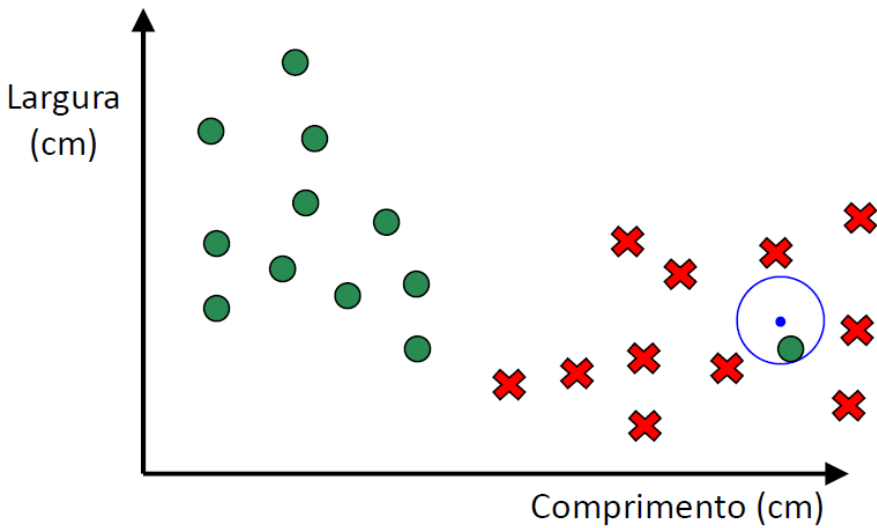
Como definir K:

Escolher um K para todo o conjunto de teste é uma tarefa complexa

A solução mais comum é testar diversos valores de K e adotar aquele que seja mais acurado

Indica-se utilizar um valor de K impar para evitar empates

KNN



Árvores de Decisão

Estratégia de aprendizado supervisionado

São construídas a partir dos valores das características (e também da classe)

São indicadas principalmente para problemas com valores discretos (especialmente atributos categóricos)

Estratégia robusta à presença de ruídos

Quanto menor a árvore, mais generalista e quanto maior, mais específica

Árvores de Decisão

Pode-se representar uma AD por um conjunto de regras “if-else”

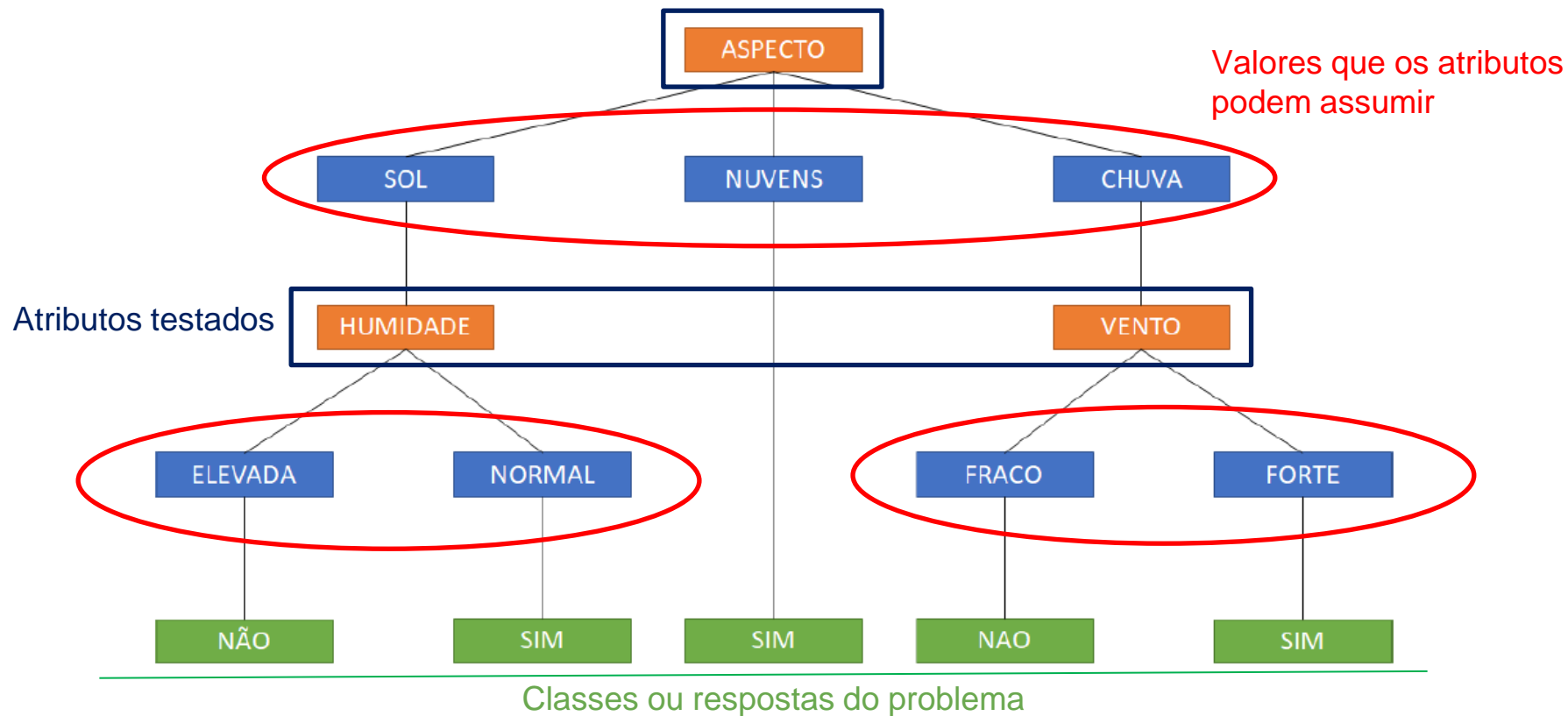
ADs classificam as instancias percorrendo a arvore da raiz até as folhas

Em cada nó, é realizado um teste sobre algum atributo da instancia

Cada ramo corresponde a um dos valores que o atributo pode assumir

Cada nó corresponde a uma regra de classificação

Árvores de Decisão



Árvores de Decisão

Construção:

Nó raíz – aquele que melhor separar o conjunto é selecionado

Um descendente do nó raiz é criado para cada valor possível deste atributo

O processo é repetido usando exemplos de treinamento associados com cada nó descendente para selecionar o melhor atributo para testar naquele nível da árvore

Processo se repete até obter um nó “puro” (somente um atributo)

Árvores de Decisão

Poda:

Quanto maior a altura, maior a especificidade

Aumenta o generalismo

Número mínimo por folhas – critério para orientar a poda da árvore

Atributos contínuos:

Identificar o ponto ideal de corte

Árvores de Decisão

AD é um modelo de fácil compreensão

Lida bem com atributos redundantes

Atributos irrelevantes podem aumentar a altura da árvore e não agregar ganho

Múltiplas árvores - Florestas aleatórias

Support Vector Machine

Estratégia de aprendizado supervisionado

Assume que as classes são linearmente separáveis

Encontra um hiperplano que oferece a separação máxima entre duas classes

Expand a entrada para um espaço de alta dimensionalidade para lidar com casos não-linearmente separáveis

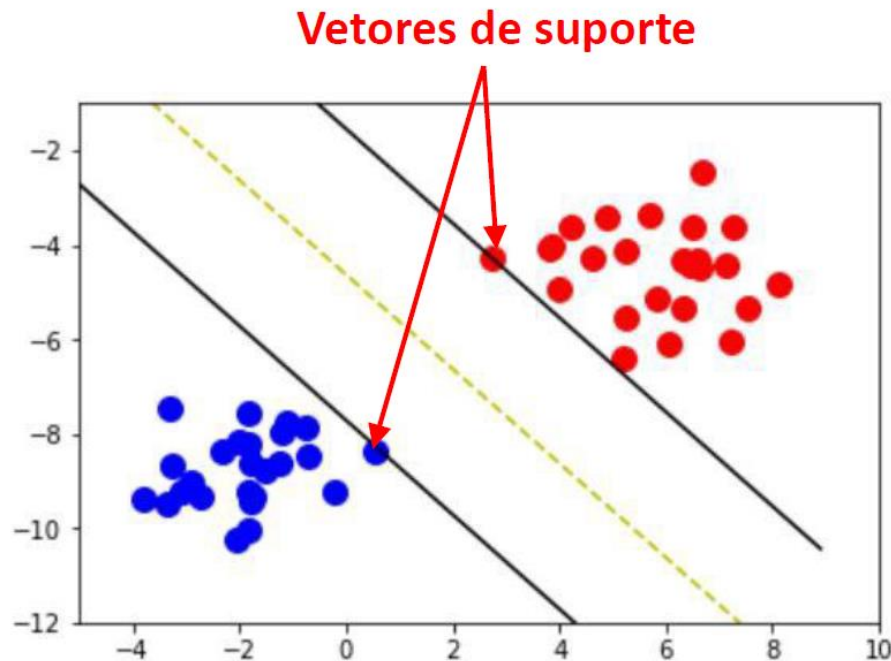
Support Vector Machine

Vetores de suporte:

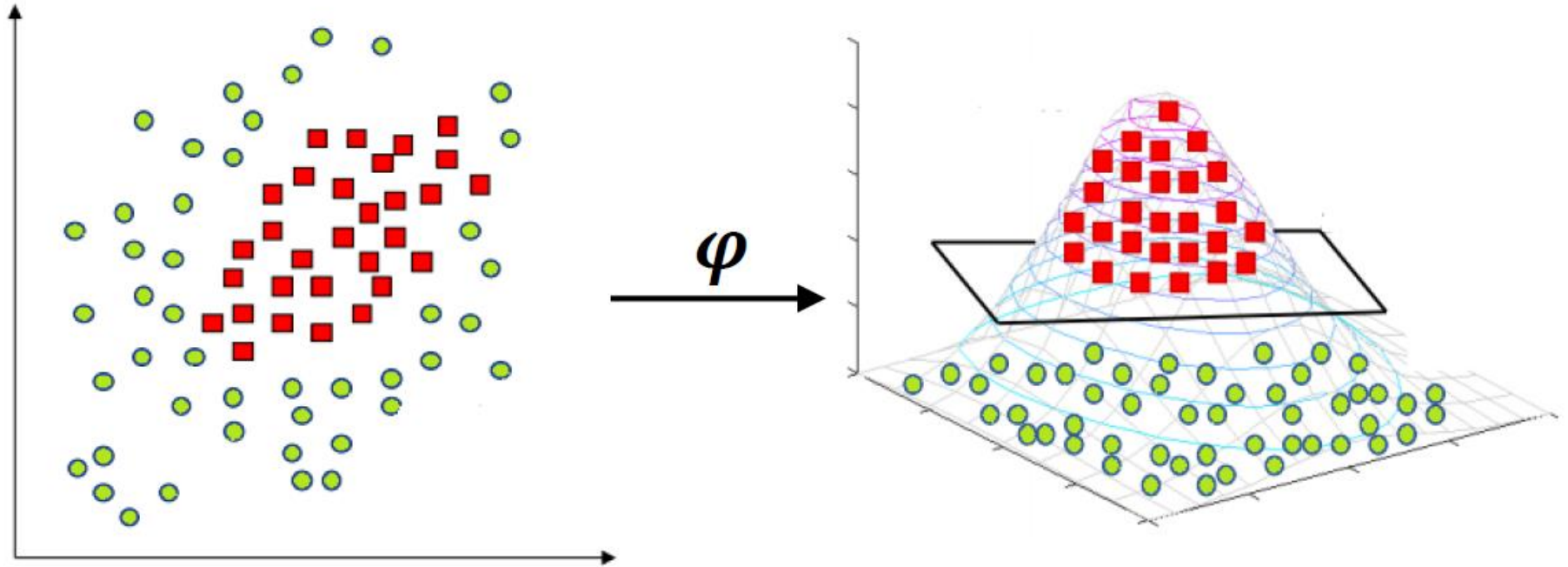
- Instancias mais próximas das fronteiras

Quando chegarem novos dados para treino, basta comparar com os vetores

Pode ser que a fronteira mude com novos dados, atualizando os vetores



Support Vector Machine



Support Vector Machine

Kernel trick

$$\mathbb{R}^d \rightarrow \mathbb{R}^X$$

Em que $d < X$

Algumas transformações de kernel:

- Gaussiano
- Polinomial
- Tangente hiperbólico
- Radial based function (RBF)

Support Vector Machine

Soft Margin

Presença de outliers

Exemplos com erros de rotulação

Permite que a margem de decisão cometa alguns poucos erros (pontos podem ficar dentro da margem ou do lado errado)

Variável de folga (C):

- Quanto maior, mais outliers serão desconsiderados
- Se $C = 0$, o problema é linearmente separável

