

INTELIGÊNCIA COMPUTACIONAL

K-MEANS: INTRODUÇÃO E CONCEITOS GERAIS

1

SUMÁRIO

– K-Means

2

Introdução

- **Clusterização:** é uma classificação sem classes predefinidas
- É como dividir meias por cor quando você não se lembra de todas as cores que você tem
- O algoritmo tenta encontrar objetos semelhantes (por características) e agrupá-los em classes
- Aqueles que têm muitas características semelhantes são unidos em uma classe

3

Aplicações

- **Sistemas de Diagnóstico:**
 - A profissão médica usa o K-Means na criação de sistemas de apoio à decisão médica mais inteligentes, especialmente no tratamento de doenças do fígado
- **Motores de Busca:**
 - A clusterização é a espinha dorsal dos motores de busca. Quando uma pesquisa é realizada, os resultados da pesquisa precisam ser agrupados e os mecanismos de pesquisa frequentemente usam clustering para fazer isso
- **Redes de sensores sem fio**
 - O algoritmo de clustering desempenha o papel de localizar os cluster heads, que coletam todos os dados em seu respectivo cluster

4

Aplicações

– Marketing e Vendas

- Personalização e targeting em marketing são um grande negócio
- Isso é conseguido examinando as características específicas de uma pessoa e compartilhando com ela as campanhas que tiveram sucesso com outras pessoas semelhantes
- **Como os clusters funcionam:** Os algoritmos de clustering são capazes de agrupar pessoas com características semelhantes e probabilidade de compra. Depois de ter os grupos, você pode executar testes em cada grupo com uma cópia de marketing diferente que o ajudará a direcionar melhor suas mensagens para eles no futuro

5

K-Means

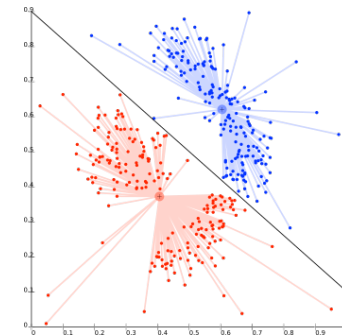
6

K-Means

- O k-Means agrupa um conjunto de n instâncias em k grupos, onde $k < n$
- Cada grupo é representado por uma centróide
- O algoritmo converge para estabilizar as centróides

7

K-Means



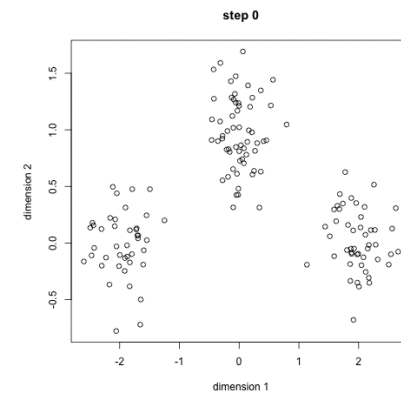
8

K-Means

1. Inicializa o conjunto de k centróides aleatoriamente
2. Assinala cada ponto do conjunto a centróide mais próxima
3. Recalcula a posição de cada centróide, i.e., centro do conjunto de pontos mais próximos
4. Volta para o passo 2, enquanto os centróides forem reposicionados

9

K-Means



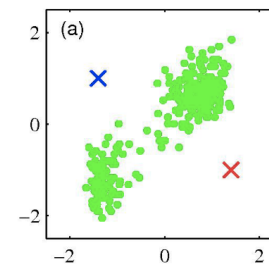
10

K-Means



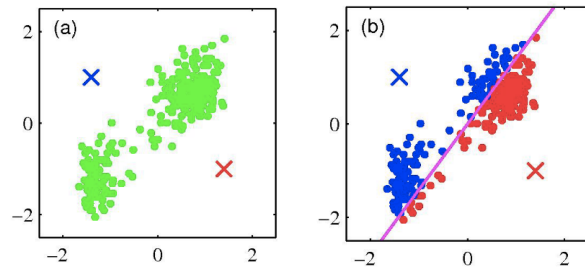
11

Exemplo



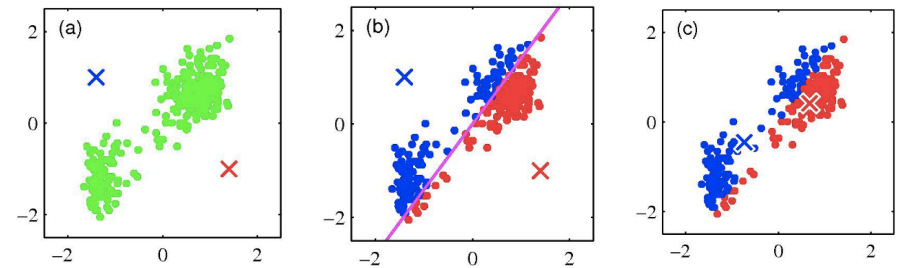
12

Exemplo



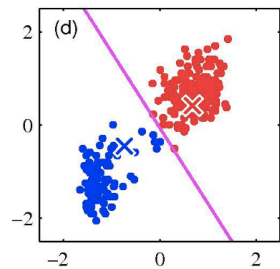
13

Exemplo



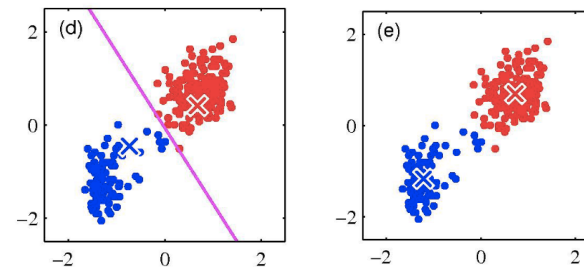
14

Exemplo



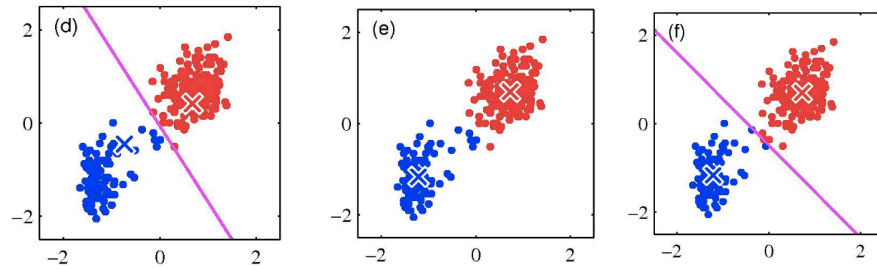
15

Exemplo



16

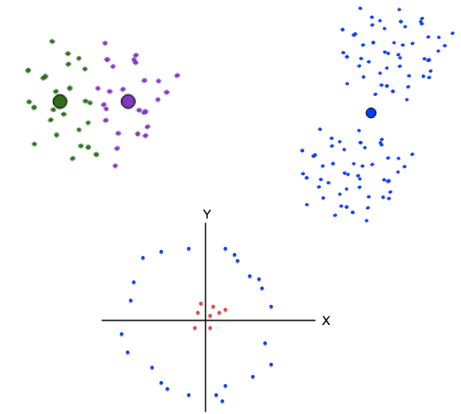
Exemplo



17

K-Means

- Qual o valor ideal de k ?
- Ótimos locais
- Alguns problemas não conseguem ser segmentados



18

REFERÊNCIAS BIBLIOGRÁFICAS

- Engelbrecht, Andries. **Computational intelligence : an introduction**. Wiley, ed. 2, 2007.
- Géron, Aurélien. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow**. O'Reilly Media.

19