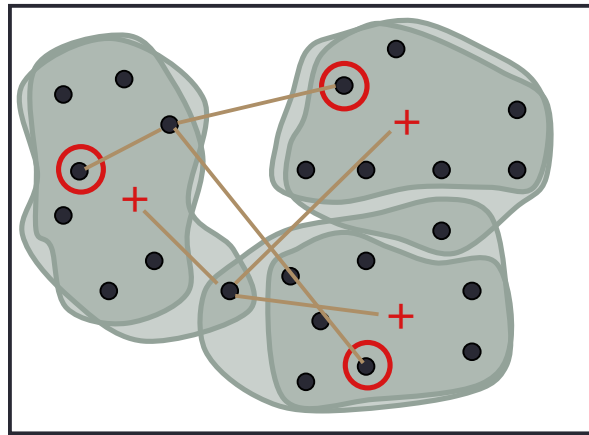


ALGORITMO K-MEANS

Cristiane Neri Nobre

Algoritmo K-means* ([MacQueen, 1967](#))

Exemplo do funcionamento do K-means para $K = 3$



1ª Iteração

*J. B. MacQueen (1967): "Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*", Berkeley, University of California Press, 1:281-297

Algoritmo K-Means

1. Selecione k pontos como centroides iniciais

Repetir

1. Forme k clusters associando cada objeto a seu centroide mais próximo
2. Recalcule o centroide de cada cluster

Até que Centroides não apresentarem mudanças

Centroide = centro de gravidade do cluster

Coordenada i = média aritmética das coordenadas i de seus objetos constituintes.

Observações

- Boa parte dos clusters já convergem nos primeiros passos do algoritmo, ficando somente uma quantidade pequena de clusters que ainda modificam.
- Assim, a condição de parada do algoritmo é substituída por "***até que somente 1% dos objetos mudam de clusters***".
- Complexidade do K-means é $O(n * K * I * d)$
 - N = número de instâncias
 - K = Número de clusters
 - I = Número de iterações
 - d = Número de atributos

Verifique outros métodos de agrupamento:

DBSCAN

https://www.youtube.com/watch?v=Xw_ig_yzHyU

<https://www.youtube.com/watch?v=s5nn4xxA2XA>

Veja esta série de vídeos:

<https://www.youtube.com/watch?v=7-vwS8Jh6eo>

Neste vídeo, aprendemos:

1. Algoritmo Kmeans

No próximo vídeo, veremos como **rodar o Kmeans no ambiente WEKA**