**Exercícios sobre Agrupamento (Clustering)**

1. Sobre o algoritmo de agrupamento K-means, assinale V (verdadeira) ou F (falsa) para cada uma das afirmações abaixo:

( ) O *K-means* (*K-média*) é um algoritmo de agrupamento estatístico que permite particionar um conjunto de dados em *K* clusters (grupos) disjuntos.

( ) Os centros iniciais (*centroids*) dos *K* clusters podem ser escolhidos aleatoriamente.

( ) O algoritmo *K-means* utiliza distância Euclidiana.

( ) *K-means* é um exemplo de técnica de aprendizagem supervisionada.

( ) Espera-se maximizar a distância entre elementos do mesmo cluster e minimizar a distância entre elementos de clusters diferentes.

( ) Um agrupamento pode ser utilizado para identificar as futuras classes de um processo de classificação.

2) Sobre o algoritmo de agrupamento K-means, assinale V (verdadeira) ou F (falsa) para cada uma das afirmações abaixo: (2.0)

( ) O número de cluster (K) é encontrado automaticamente.

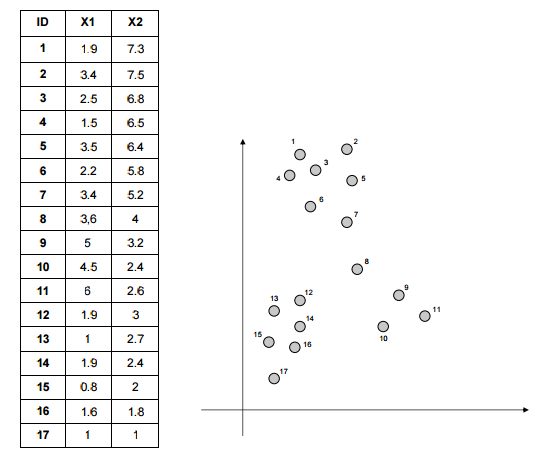
( ) Não funciona para dados contínuos.

( ) Pode utilizar outras distâncias além da Euclidiana.

( ) É um exemplo de técnica de aprendizagem não supervisionada.

( ) Espera-se minimizar a distância entre elementos do mesmo cluster e maximizar a distância entre elementos de clusters diferentes.

3) Considere o seguinte conjunto de exemplos não rotulados:



Use o WEKA, considere K=3 e responda:

1. Qual será o valor de cada *centróide*?
2. Como serão agrupados os pontos? Mostre no gráfico.