

2ª Avaliação

Questão 1: (1 pt)

A Figura 6.4 representa o resultado de um algoritmo de agrupamento, onde os elementos pontilhados não foram agrupados. Baseado nesta figura, qual das afirmações abaixo podemos afirmar ser verdadeira?

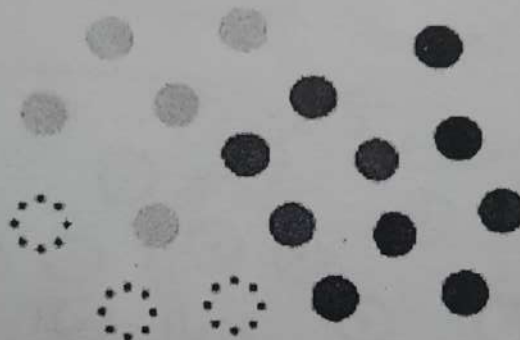


Figura 6.4

- a) Foi utilizado um algoritmo K-means.
- b) Foi utilizado um algoritmo K-medoids.
- c) Trata-se de um agrupador do tipo hierárquico.
- ☒ d) Trata-se de um agrupador por densidade DBSCAN.

Questão 2: (1 pt)

O cluster na Figura 6.5 é um dendrograma. Baseado nesta figura, qual das afirmações abaixo podemos afirmar ser verdadeira?

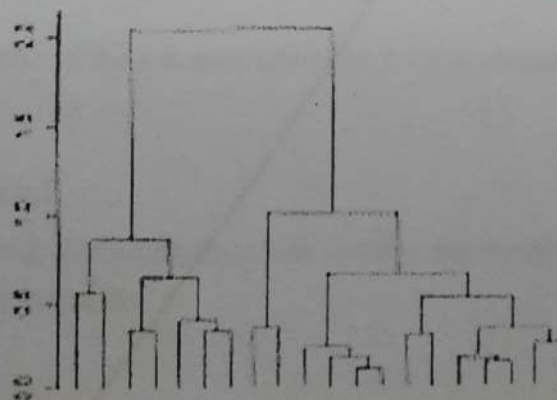


Figura 6.5

- a) Foi utilizado um algoritmo K-means.
- b) Foi utilizado um algoritmo K-medoids.
- ☒ c) Trata-se de um agrupador do tipo hierárquico.
- d) Trata-se de um agrupador por densidade DBSCAN.

Questão 3: (1 pt)

Observe os agrupamentos A e B na Figura 6.6. Sobre estes agrupamentos, escolha a alternativa correta.

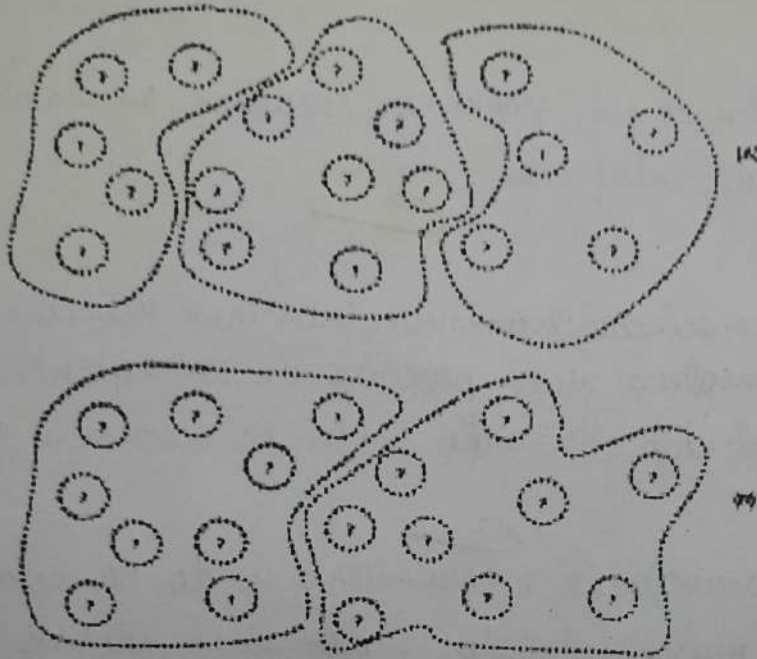


Figura 6.6: Agrupamento, A e B

- ☒ Os agrupamentos podem ter sido gerados pelo mesmo algoritmo, porém A com  $K=3$  e B com  $K=2$ .
- b) Os elementos de A são diferentes de B, por isso os grupos gerados são diferentes.
- c) O fato de não haver ruído demonstra que foi utilizado DBSCAN.

Questão 4: (1 pt)

Explique sucintamente em que consiste o aprendizado de máquina não-supervisionado.

Questão 5: (2 pts)

Faça uma sucinta distinção entre as tarefas de mineração: Agrupamento de dados e Extração de regras de associação.

Questão 6: (2 pts)

Comente sobre pelo menos uma aplicação prática da Extração de regras de associação.

Questão 7: (2 pts)

Com relação a Extração de regras de associação, explique em que consistem os parâmetros de suporte e confiança.



Q4:

R- No aprendizado de Máquina não supervisionado os dados não ~~estão~~ <sup>são</sup> rotulados, ~~de~~ ~~modo~~ ~~que~~ desse modo a máquina não recebe nenhuma "ajuda" direta, com o objetivo de obrigá-la a entender e processar os dados sozinha, um exemplo desse tipo de aprendizado é a tarefa de agrupamento de dados.

Q5:

R- No agrupamento de dados buscamos organizar/separar as instâncias em grupos a partir da similaridade entre elas.

Na extração de regras por associação buscamos encontrar relações do tipo: um conjunto de antecedentes implica num conjunto de consequentes, isto é, a ocorrência dos antecedentes obriga, até certo ponto, na ocorrência do consequente

Q6:

R- A extração de regras por associação é amplamente usada na área de comércio, pois a partir das regras obtidas é possível adotar estratégias, como mudança de produtos ou organização do mesmo, visando o aumento do consumo e consumo.

Q7:

R- O parâmetro de suporte de uma relação  $(X \rightarrow Y)$  expressa a frequência com que ela ocorre no dataset para assim podermos classificá-la como frequente quando ela for maior ou igual a um suporte mínimo. É dada por

$$\text{Suporte}(X \rightarrow Y) = \frac{|X \cup Y|}{\text{Total}}$$

O parâmetro de confiança expressa a validade de uma relação quantificando o quanto ela é ~~valida~~ válida a partir de um valor de confiança mínima. É dada por:

$$\text{Confiança}(X \rightarrow Y) = \frac{|X \cup Y|}{|X|}$$