

Teoria T06

Preencha o círculo dos itens verdadeiros e deixe em branco o dos falsos.

- O algoritmo K-Means busca minimizar a soma das distâncias intra-classe para um dado número de clusters.
- Se utilizarmos o método *elbow*, o algoritmo K-Means pode convergir para o melhor K ao longo das iterações, eventualmente dispensando uma definição explícita do K inicial.
- Em um problema de agrupamento, o fato de todas as classes (com rótulos ocultos, naturalmente) possuírem médias iguais para cada característica necessariamente prejudica o desempenho do K-Means,
- O resultado do agrupamento via algoritmo K-Means++ é usualmente melhor que o do K-Means por não utilizar aleatoriedade para escolha de nenhum dos centroides iniciais.
- Em detrimento da melhor qualidade do agrupamento, o K-Means++ possui tempo de processamento usualmente mais alto que o do K-Means original em virtude do overhead para escolha dos centroides iniciais.
- Coeficiente Silhoutte com valor próximo à 1 (um) significa ótimo resultado de agrupamento
- Coeficiente Silhouette com valor bem próximo a zero significa que a maioria dos dados foram agrupados em clusters errados.
- No DBSCAN, fixando-se o número mínimo de pontos necessários numa região para que um *core-point* deseje definido, se aumentarmos o *eps*, a tendência é encontrarmos menos clusters.
- No DBSCAN, fixando-se o *eps*, se aumentarmos o *minPts*, a tendência é que um número maior de dados sejam considerados ruídos.
- Em um problema de agrupamento, o fato de todas as classes (com rótulos ocultos, naturalmente) possuírem médias iguais para cada característica necessariamente prejudica o desempenho do DBSCAN.