

# Agrupamento (Clustering)

Revisar todo o material

1.1. - Business Understanding- Explicar melhor a escolha do tema como o porquê e quais suas utilidades aproximadamente 2 páginas.

- Quais as Tarefas de Data Science que são elegíveis ?

1.2 – Data Understanding e Data Preparation:

- Descrever Variáveis (Features) do problema (nomes, tipos, domínio de valores).
- Foi preciso conversão de atributo Numérico para Nominal (ou vice-versa), conversão de Real para Inteiro, normalização, dentre outros. ?

1.3 – Modeling e Evaluation

- Construir uma tabela com “Tipos de Inicialização (Randômica e Método K-Means++)” e “Medidas de Distância (Euclidiana e Manhattan)”
- Construir uma tabela, com as técnicas de testes (“Percentage Split” e “Cross Validation”).
- Evaluation: Comente os resultados obtidos com base no critério “WCSSE”, ou seja, Qualidade do Modelo.
- Comente os resultados obtidos com base no critério Qualidade dos Clusters. Explicar o significado de cada cluster formado!
- Por fim, qual o melhor Modelo obtido? Justifique!

Compreender o funcionamento de cada biblioteca, cada chamada de função, explicar os gráficos.

Criar uma documentação de explicação de funcionalidade de modo a deixar o código mais limpo.

Se possível traduzir o código e CSV para tornar mais clara a compreensão.

Recortar e colar prints do código.