

Trabalho de Avaliação: Comparação de Algoritmos de Agrupamento

Objetivo Geral:

Implementar e comparar três algoritmos de agrupamento não supervisionado: Kohonen (SOM), K-means e DBSCAN, utilizando o mesmo conjunto de dados (Iris Dataset), sem utilizar bibliotecas que implementem os algoritmos prontos.

Regras Gerais:

1. Máximo de 3 alunos por grupo;
2. Proibido o uso de bibliotecas que já implementem os algoritmos, como *sklearn.cluster.KMeans*, *sklearn.cluster.DBSCAN*, etc.;
3. Permitido o uso de bibliotecas auxiliares como numpy, matplotlib e pandas
4. Cada algoritmo deve ser implementado do zero
5. O dataset "Íris" pode ser carregado via *sklearn.datasets.load_iris* ou carregado por outro método a escolha do aluno;
6. Cada grupo deve testar e discutir o impacto da variação de parâmetros equivalentes nos três algoritmos.

Parâmetros a serem investigados

Cada grupo deverá realizar experimentos comparando **diferentes configurações** dos seguintes parâmetros, que possuem papéis equivalentes nos três métodos:

Conceito	K-means	DBSCAN	Kohonen (SOM)
Número de grupos esperados	k (número de clusters)	N/A diretamente ¹	Tamanho da grade (ex: 2x2, 3x3 etc.)
Distância mínima entre pontos	Implícita pelo centróide	eps (raio de vizinhança)	Raio de vizinhança (σ)
Tamanho mínimo do grupo	N/A	minPts	N/A
Critério de parada	Mudança mínima dos centros	Todos os pontos visitados	Número de épocas (epochs)
Taxa de aprendizagem	N/A	N/A	learning_rate

¹ No caso do DBSCAN, o número de grupos é emergente. A análise deve discutir essa diferença e como ela afeta os resultados.

Entregáveis:

1. **Código-fonte** com as três implementações
2. **Relatório em PDF**, no formato IEEE, com:
 - Introdução ao problema de agrupamento
 - Descrição conceitual e matemática dos três algoritmos
 - Explicação do dataset Iris
 - Detalhes das implementações (incluir o link do github com os códigos)
 - Análise dos parâmetros testados (com gráficos)
 - Visualizações dos agrupamentos para diferentes valores de parâmetros
 - Discussão crítica dos resultados e impacto dos parâmetros
 - Conclusão final

Critérios de Avaliação (Total: 10 pontos):

Critério	Pontos
Implementação correta do K-means	1,5
Implementação correta do DBSCAN	2,0
Implementação correta do SOM	1,5
Análise dos parâmetros e sua influência nos resultados	2,5
Clareza e profundidade do relatório	2,5

Entrega

Enviar o pdf para o e-mail pablorange@cos.ufrj.br, com cópia para cmicelifarias@cos.ufrj.br.

- No assunto, colocar “Entrega do Trabalho GESCON 2025.1”.
- No corpo do e-mail, listar o DRE e o nome dos alunos.

Prazo

07 de Julho de 2025.