



Como funciona o aprendizado não-supervisionado

☰ Ciclo	Ciclo 05: Aprendizado não-supervisionado
# Aula	33
🕒 Created	@February 24, 2023 8:50 AM
☑ Done	☑
☑ Ready	☑

Objetivo da Aula:

- ☐ Aprendizado não-supervisionado
- ☐ Tarefas de Clusterização
- ☐ Exemplos de algoritmos de Clusterização
- ☐ Resumo
- ☐ Próxima aula

Conteúdo:

▼ 1. Aprendizado não-supervisionado

No aprendizado não-supervisionado, temos um conjunto de dados de treinamento com característica observadas do fenômeno, mas não temos uma variável numérica ou um rótulo de classificação das observações.

Nesse tipo de aprendizado, nosso único interesse é descobrir padrões.

▼ 1.1 Descobrir padrões nos dados

Padrões são comportamentos ou características similares de pessoas ou produtos dentro de um determinado grupo.

▼ 1.1.1 Exemplo

Pessoas entre 24 e 35 anos, residentes na região Sudeste, com ensino superior completo na área de exatas, constituinte de famílias com até 4 membros, praticantes de exercício físico duas vezes por semana, que ingerem 2L de água por dia e moram em casas, comprar eletrônicos acima de R\$ 5.000 em média.

▼ 2. Tarefas de Clusterização

A técnica de clusterização agrupa objetos similares em grupos que apresentam padrões de comportamento e/ou características parecidas.

Esses grupos de objetos similares entre si formam padrões que são usados para entender melhor estruturas de dados, segmentação de clientes por características ou comportamentos similares.

▼ Exemplos:

1. **Segmentação de clientes:** usando dados de transações, a clusterização pode identificar grupos de clientes com comportamentos de compra semelhantes, como frequência de compra, valor de compra ou produtos comprados. Esses grupos podem ser usados para personalizar ofertas, campanhas de marketing ou para entender melhor as preferências dos clientes.
2. **Identificação de fraudes:** a clusterização pode ser usada para detectar atividades suspeitas ou fraudes em transações financeiras. Por exemplo, grupos de transações atípicas ou fora do padrão podem ser identificados como potenciais fraudes.

3. **Segmentação de imagens:** a clusterização pode ser usada para segmentar imagens em regiões com características similares. Por exemplo, em imagens de satélite, a clusterização pode ser usada para segmentar áreas com características similares, como uso do solo, cobertura vegetal ou densidade populacional.
4. **Análise de textos:** a clusterização pode ser usada para agrupar textos similares com base em tópicos ou temas. Isso pode ser usado para análise de sentimentos, identificação de tendências ou para entender melhor a opinião pública sobre determinado assunto.

▼ 3. Exemplos de algoritmos de Clusterização

▼ 3.1 Clusterização

1. K-Means
2. Hierarchical Clustering Analysis (HCA)
3. DBSCAN
4. Meanshift
5. Gaussian Mixture Model (GMM)

▼ 4. Resumo

1. O aprendizado não-supervisionado tem o objetivo de agrupar indivíduos com características ou comportamentos semelhantes, para encontrar padrões.

▼ 5. Próxima aula

K-Means - Teoria