

Modelos de Clusterização: descoberta de padrões e segmentação de dados

Curso: Modelos de Clusterização: descoberta de padrões e segmentação de dados
Carga Horária: 20 horas (12 horas ao vivo + 8 horas atividades orientadas)
Categorias: Ciência de Dados, Machine Learning
Nível: Intermediário
Plataforma: Open Source
Especialista: A definir
Formato: Ao vivo

Sobre o curso

O curso Modelos de Clusterização — Descoberta de Padrões e Segmentação de Dados foi desenvolvido para profissionais de dados que atuam com análise exploratória, segmentação e descoberta de padrões, em contextos em que não há rótulos previamente definidos.

Ao longo do curso, o participante aprenderá a aplicar técnicas de clusterização para entender a estrutura dos dados, identificar grupos semelhantes e gerar hipóteses analíticas que auxiliem investigações mais profundas. O conteúdo enfatiza a interpretação crítica dos agrupamentos, o impacto das decisões de pré-processamento e os cuidados necessários para evitar segmentações artificiais ou enganosas.

O curso prepara o profissional para utilizar clusterização como instrumento analítico, apoiando estudos exploratórios, análises comportamentais e entendimento de bases complexas no dia a dia do trabalho.

Módulos do Curso

Módulo 1 — Fundamentos de Clusterização

— Aprendizado não supervisionado

- Similaridade e distância (Euclidiana e Cosseno)
- Aplicações de clusterização

Módulo 2 — Algoritmos de Clusterização

- K-Means
- Hierárquicos
- DBSCAN

Módulo 3 — Avaliação de Clusters

- Métodos de validação interna
- Silhuete Score
- Davies Bouldin
- Índice de Jaccard
- Calinski Harabasz
- Limitações da avaliação

Módulo 4 — Pré-processamento e Escalonamento

- Normalização e padronização
- Impacto das variáveis
- Redução de dimensionalidade
- Engenharia de Features

Módulo 5 — Interpretação e Uso Analítico

- Segmentação de dados
- Comunicação de resultados
- Comunicando grupos ao comitê executivo
- Riscos de interpretação indevida

O que você vai aprender?

- Compreensão dos fundamentos de clusterização
- Aplicação de algoritmos não supervisionados
- Avaliação e interpretação de agrupamentos
- Uso crítico de clusters em análises exploratórias

Pra quem é o curso?

Este curso é indicado para:

- Cientistas e Analistas de Dados
- Profissionais que trabalham com segmentação
- Pessoas interessadas em aprendizado não supervisionado

Pré-requisitos

- Estatística básica
- Manipulação de dados
- Conceitos introdutórios de Machine Learning

rox