

10 - CLUSTERING II

Aprendizagem 2024/2025

EM

- Algoritmo baseado em modelo, onde os clusters são representados por distribuições
- Composto por dois passos essenciais (expectation e maximization)

E - STEP

- Atribuir pontos a clusters
- Para cada ponto, calcular o seguinte para cada cluster:
 - (a) Prior (dado)
 - (b) Likelihood (ponto na distribuição)
 - (c) Joint probability (multiplicação de (a) com (b))
- Normalizar os valores para cada ponto

M - STEP

- Recalcular parâmetros (geralmente médias e variâncias para 1 dimensão, ou médias e matrizes de covariâncias para mais que 1 dimensão)

SUMÁRIO

- Ficha 10: 4, 5, 6