BÁO CÁO BÀI TẬP LỚP KHAI PHÁ DỮ LIỆU

# Tổng quan về phân cụm dữ liệu

## Tổng quan

Phân cụm dữ liệu là pPhân tích cụm hay đơn giản là phân cụm là quá trình phân vùng một tập hợp các đối tượng dữ liệu (hoặc các quan sát) thành các tập con. Mỗi tập hợp con là một cụm, sao cho các đối tượng trong một cụm tương tự với nhau, nhưng không giống với các đối tượng trong các cụm khác.

Tập hợp các cụm kết quả từ phân tích cụm có thể được gọi là phân cụm.Các phương pháp phân cụm khác nhau có thể tạo ra các chuỗi khác nhau trên cùng một tập dữ liệu. Các phân vùng không được thực hiện bởi con người, nhưng bởi thuật toán phân cụm. Do đó, phân cụm hữu ích ở chỗ nó có thể dẫn đến việc khám phá các nhóm chưa biết trước đây trong dữ liệu.

## abc

# Thuật toán K-Medoid

# Triển khai thuật toán