

Característica	K-Means	GMM (Gaussian Mixture Models)
Concepto Principal	Agrupa puntos de datos en clusters basándose en la distancia a los centroides. Cada punto pertenece exclusivamente a un cluster.	Modela los datos como una mezcla de distribuciones gaussianas. Un punto puede pertenecer a múltiples clusters con diferentes probabilidades.
Forma de los Clusters	Asume que los clusters son esféricos y de tamaño similar.	Es flexible y puede manejar clusters con formas elípticas o irregulares.
Enfoque	Particional y determinista. Asigna cada punto al cluster más cercano.	Probabilístico y basado en modelos. Estima la probabilidad de que cada punto pertenezca a cada cluster.
Ventajas	- Simple y rápido de implementar.- Ideal para grandes conjuntos de datos.- Fácil de interpretar.	- Más flexible para clusters complejos.- Proporciona información probabilística sobre la pertenencia a los clusters.- Robusto ante la superposición de clusters.
Desventajas	- Sensible a la inicialización de los centroides.- No es ideal para clusters de formas irregulares.- Sensible a los valores atípicos (outliers).	- Más complejo y computacionalmente intensivo. - Requiere más datos para un rendimiento óptimo. - La interpretación puede ser menos directa.
Algoritmo Clave	Algoritmo iterativo que busca minimizar la suma de los cuadrados de las distancias dentro de los clusters.	Algoritmo de Maximización de Expectativa (EM), que ajusta iterativamente los parámetros de las distribuciones gaussianas.
Aplicaciones Típicas	- Segmentación de clientes en clusters bien definidos.- Compresión de imágenes.- Clasificación de documentos simples.	- Detección de anomalías.- Análisis de imágenes y señales complejas.- Modelado de distribuciones de datos más sofisticadas.

