

Característica	Descripción
Concepto Principal	Es un modelo de clustering probabilístico que asume que los datos están compuestos por una mezcla de varias distribuciones gaussianas (o normales) con parámetros desconocidos.
Cómo Funciona	A diferencia de K-Means, no asigna cada punto a un solo cluster. En su lugar, calcula la probabilidad de que un punto de datos pertenezca a cada uno de los clusters, lo que permite una asignación "suave" o "borrosa" .
Algoritmo Clave	Utiliza el algoritmo de Expectation-Maximization (EM) para estimar de forma iterativa los parámetros de cada distribución gaussiana (la media, la covarianza y el peso de cada cluster).
Ventajas	<ul style="list-style-type: none"> <li>- Flexibilidad: Puede modelar clusters con formas elípticas y tamaños variados, no solo esféricos como K-Means.</li> <li>- Información probabilística: Proporciona un nivel de confianza (probabilidad) para la pertenencia de cada punto a un cluster.</li> <li>- Detección de anomalías: Los puntos con una densidad de probabilidad muy baja pueden ser identificados como anomalías.</li> </ul>
Desafíos	<ul style="list-style-type: none"> <li>- Complejidad: Es más complejo y computacionalmente intensivo que K-Means.</li> <li>- Elección de parámetros: Requiere decidir el número de componentes (clusters) y el tipo de matriz de covarianza, lo cual a menudo se hace usando métricas como AIC o BIC.</li> </ul>
Aplicaciones	<ul style="list-style-type: none"> <li>- Análisis de imágenes y procesamiento de señales.</li> <li>- Detección de anomalías en datos.</li> <li>- Segmentación de mercado cuando los grupos tienen formas irregulares.</li> </ul>