

Modulo 2

LECCIÓN 4

Algoritmos de Agrupamiento

Algoritmos de Agrupamiento

Tiempo de ejecución: 8 horas

Objetivos de la Lección:

- Comprender los conceptos fundamentales de los algoritmos de agrupamiento.
- Familiarizarse con diferentes técnicas de agrupamiento, incluyendo K-Means, Agrupamiento Jerárquico, DBSCAN, Mean Shift, Gaussian Mixture Models (GMM) y Spectral Clustering.
- Aprender a aplicar y evaluar estos algoritmos utilizando conjuntos de datos de ejemplo.

Contenido de la Sesión:

1. Introducción a los Algoritmos de Agrupamiento

- Definición de agrupamiento y su importancia en el aprendizaje no supervisado.
- Conceptos básicos como centroides, distancia euclidiana y función de similitud.
- Aplicaciones de los algoritmos de agrupamiento en diferentes campos.

2. K-Means

- Explicación detallada del algoritmo K-Means.
- Proceso de agrupamiento utilizando K-Means.
- Consideraciones sobre la inicialización de centroides y selección del número de clusters.
- Ejemplos prácticos de aplicación de K-Means en Python.

3. Agrupamiento Jerárquico

- Concepto de agrupamiento jerárquico y sus dos enfoques: aglomerativo y divisivo.
- Métodos de enlace (single-linkage, complete-linkage, average-linkage).
- Visualización de dendrogramas y selección del número de clusters.
- Implementación de agrupamiento jerárquico en Python.

4. DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

- Principio de funcionamiento del algoritmo DBSCAN.
- Definición de parámetros como epsilon y minPts.
- Identificación de puntos centrales, puntos de frontera y puntos de ruido.
- Ejemplos de aplicación de DBSCAN en la detección de anomalías y clustering de datos.

5. Mean Shift

- Descripción del algoritmo Mean Shift y su aplicación en la estimación de densidad.
- Concepto de ventana de búsqueda y convergencia hacia los modos de densidad.
- Implementación de Mean Shift en Python y ejemplos de uso en agrupamiento de datos.

6. Gaussian Mixture Models (GMM) y Spectral Clustering Introducción a los modelos de mezclas gaussianas (GMM) y su relación con el agrupamiento.

- Concepto de clustering espectral y su aplicación en grafos.
- Comparación entre GMM y Spectral Clustering.
- Ejemplos prácticos de aplicación de ambos algoritmos en Python.

Evaluación:

- Evaluación continua a través de ejercicios prácticos durante la lección.
- Evaluación final mediante la implementación de al menos dos de los algoritmos aprendidos en un conjunto de datos real y la interpretación de los resultados obtenidos.

Recursos:

- Conjuntos de datos de ejemplo para la práctica.
- Entorno de programación en Python con bibliotecas como scikit-learn y matplotlib.
- Material didáctico con explicaciones detalladas y ejemplos de código para cada algoritmo de agrupamiento.

