

INTRODUCCIÓN

- **Tema:** Métricas de modelos de agrupamiento
- **Resultado de Aprendizaje:** Explicar, calcular y analizar métricas de calidad de partición como diámetro y separación en modelos de agrupamiento.

1. LECCIÓN EN CASA

1.1 Adquisición de concepto

Para la adquisición del concepto, se solicitará al estudiante interactuar con ChatGPT mediante la siguiente serie de indicaciones:

1. Interactuar con ChatGPT mediante los siguientes *prompts*, leyendo detenidamente el *prompt* y su respuesta:

Prompt 1. Vas a ser mi profesor de la asignatura de Aprendizaje Automático, te daré instrucciones y me explicarás de manera clara y formal lo que te pida. Quiero que seas muy preciso con los conceptos matemáticos, pero también que uses ejemplos simples para ilustrar los conceptos si es necesario. Sé ameno y paciente. ¿Entendido?

Prompt 2. ¿Qué son las métricas de calidad de partición en un modelo de agrupamiento? Explícalo en términos generales sin entrar en detalles matemáticos aún.

Prompt 3. Explícame qué es el diámetro de un grupo en un modelo de agrupamiento. Usa una definición matemática simple.

Prompt 4. Ahora explícame qué significa la separación entre dos grupos. Define este concepto de manera matemática.

Prompt 5. ¿Por qué son importantes el diámetro y la separación para evaluar un modelo de agrupamiento? Dime cómo se relacionan con la calidad del agrupamiento.

Prompt 6. Dame un ejemplo general de cómo calcular el diámetro de un grupo y la separación entre dos grupos usando conceptos teóricos. No utilices datos numéricos todavía.

Prompt 7. Ahora sí, dame un ejemplo numérico donde calcules el diámetro y la separación usando datos de puntos en un espacio bidimensional.

2. Visualiza el siguiente video: ¿Qué tan buenos son tus Clusters?.
3. Continúa la interacción con ChatGPT con las preguntas sobre el video que acabas de ver.
4. Realiza el cuestionario del aula virtual.

1.2 Personalización de la actividad

Se la consigue en la interacción personal de cada estudiante con ChatGPT, solicitando también continuar la interacción hasta que se sienta preparado en el tema.

1.3 Solventación de dudas

En caso de tener dudas sobre el tema, se solicitará al estudiante interactuar con ChatGPT para solventarlas.

1.4 Micro-tarea

Para realizar un seguimiento de la actividad, se solicitará al estudiante copiar el enlace del chat como evidencia del proceso. Adicionalmente, se le pedirá realizar el cuestionario del aula virtual. El cuestionario se encuentra detallado en el Anexo.

2. TAREAS EN CLASE

2.1 Visión conjunta

Se muestra la relación entre las actividades realizadas en casa y las tareas a realizar en clase. De manera específica, cómo el cálculo de determinantes se relaciona con sus diferentes propiedades y con la inversión de matrices.

2.2 Retroalimentación

Se brinda retroalimentación a los estudiantes sobre las respuestas dadas en la micro-tarea.

2.3 Actividad de aplicación

Se utilizará el siguiente código base para calcular métricas en modelos de agrupamiento: 06-Evaluacion-de-Modelos-II.ipynb

2.4 Micro-evaluación

No se realizará micro-evaluación en esta clase.

ANEXO

2.5 Métrica Cluster

1. MétricaCluster01

¿Qué es un cluster en el contexto de análisis de datos?

- a) Un punto de datos que se utiliza para realizar cálculos estadísticos.
- b) Un agrupamiento natural de datos que comparten similitudes. (100 %)
- c) Un algoritmo utilizado para clasificar datos en categorías.
- d) Una métrica para evaluar la calidad de una visualización.

2. MetricaCluster02

¿Qué mide el coeficiente de silueta en el análisis de clusters?

- a) El número total de clusters en los datos.
- b) Qué tan compactos están los puntos dentro de un cluster y qué tan separados están los clusters entre sí. (100 %)
- c) La distancia promedio entre todos los puntos del dataset.
- d) El rendimiento computacional del algoritmo utilizado.

3. MetricaCluster03

¿Cuál de las siguientes afirmaciones describe correctamente el índice Davies-Bouldin?

- a) Es una métrica que evalúa la separación de los clusters únicamente.
- b) Es una métrica que mide la cohesión y separación de los clusters, donde valores cercanos a 0 indican una mejor clusterización. (100 %)
- c) Evalúa exclusivamente la cantidad de datos en cada cluster.
- d) Se utiliza para calcular el número óptimo de clusters en el método K-Medias.

4. MetricaCluster04

¿Cuál es la principal diferencia entre los algoritmos K-Medias y Propagación de Afinidad?

- a) K-Medias requiere especificar el número de clusters como entrada, mientras que Propagación de Afinidad los determina automáticamente. (100 %)
- b) K-Medias clasifica datos basándose en probabilidades, mientras que Propagación de Afinidad utiliza métricas geométricas.
- c) Propagación de Afinidad requiere más memoria que K-Medias para procesar grandes datasets.
- d) K-Medias funciona solo con datos numéricos, mientras que Propagación de Afinidad trabaja con datos categóricos.

2.6 Clase Invertida**1. ClaseInvertida-Chat**

Copia el enlace del chat con ChatGPT como evidencia de la actividad realizada en casa.

.....
.....

Información para evaluadores:

- Acceder al enlace.

2. ClaseInvertida-Dudas

¿Qué dudas tuviste sobre el tema cubierto en la actividad?

.....

.....

Información para evaluadores:

- Solo para registro.