



Práctica: Descubriendo patrones con K-means



Objetivo general:

Aplicar el algoritmo de K-means para identificar agrupamientos naturales en un conjunto de datos, analizar su comportamiento y reflexionar sobre su utilidad práctica en el análisis de datos.



Modalidad de trabajo

Los estudiantes trabajarán en grupos de 2 a 3 personas, como lo han venido haciendo en las misiones anteriores. El trabajo es colaborativo, por lo que se espera que todos los integrantes participen activamente en el desarrollo, análisis y presentación de la actividad.



Recursos necesarios

- Jupyter Notebook o Google Colab.
- Conexión a internet.
- Librerías: numpy, matplotlib, pandas, sklearn.
- Acceso al artículo:

<https://www.aprendemachinelearning.com/k-means-en-python-paso-a-paso/>



Pasos de la actividad


1. Exploración inicial (15 minutos)

Visiten y lean en grupo el artículo del enlace anterior.




Conversen y respondan:

- ¿Qué es el algoritmo K-means?
- ¿Qué problema resuelve?
- ¿Cómo funciona el proceso de agrupamiento?
- ¿Por qué es importante definir correctamente el número de clusters?


 Cada grupo debe escribir un breve resumen (3 a 5 líneas) de su comprensión inicial.

2. Implementación práctica (40 minutos)

Sigan paso a paso el procedimiento descrito en el artículo. Pueden utilizar directamente el conjunto de datos propuesto allí, generado con `make_blobs`.

 Pasos sugeridos:

1. Importar las librerías necesarias
2. Generar o reutilizar los datos del artículo
3. Visualizar los datos originales
4. Aplicar el algoritmo K-means con un número inicial de clusters
5. Visualizar los resultados obtenidos con los grupos coloreados
6. Aplicar el método del codo para determinar el número óptimo de clusters
7. Volver a aplicar K-means con el número óptimo y observar los resultados

 Sugerencia: pueden modificar los parámetros de los datos (como la cantidad de puntos o el número de características) para explorar variaciones en el comportamiento del algoritmo.

3. Análisis e interpretación (20 minutos)


Luego de realizar la práctica, cada grupo deberá reflexionar y responder:



- ¿Cómo se comportó el algoritmo al momento de agrupar?
- ¿Se identificaron agrupamientos claros?
- ¿Qué sucedió al cambiar la cantidad de clusters?
- ¿Consideran que el algoritmo fue eficaz en este caso? ¿Por qué?
- ¿En qué tipo de problemas reales consideran que se podría aplicar este tipo de técnica?

Estas respuestas deben incluirse en el informe final, junto con las visualizaciones de los datos.

4. Presentación final (15 minutos)

 Cada grupo dispondrá de 3 a 5 minutos para presentar:

- Una síntesis de lo que aprendieron sobre K-means
- Las decisiones que tomaron durante la práctica

- Los resultados obtenidos con sus gráficas
- Una posible aplicación del algoritmo en un contexto real (educativo, empresarial, científico, etc.)

Entrega del informe grupal

Cada grupo entregará un documento (formato PDF) que incluya:

1. Objetivo de la práctica
2. Desarrollo del procedimiento
3. Gráficos obtenidos
4. Análisis de resultados
5. Reflexión final sobre la utilidad del algoritmo