CORPORACIÓN UNIVERSITARIA IBEROAMERICANA

INGENIERÍA EN DESARROLLO DE SOFTWARE

Actividad 4 - Métodos de aprendizaje no supervisado

Jhonny Avila Quintero

Profesor: Jorge Castañeda

Inteligencia Artificial

Bogotá D.C

# Análisis de Transporte Masivo con Técnicas de Agrupamiento (Aprendizaje No Supervisado)

# Introducción

El presente informe tiene como objetivo aplicar técnicas de aprendizaje automático no supervisado, específicamente métodos de agrupamiento, para analizar un conjunto de datos simulados relacionados con el uso del transporte masivo. Este ejercicio se enmarca dentro del estudio de inteligencia artificial y busca identificar patrones o comportamientos comunes entre los usuarios del sistema de transporte. La técnica empleada fue el algoritmo K-Means, una herramienta efectiva para clasificar grandes volúmenes de datos sin etiquetar.

# Descripción del Dataset

El dataset simulado contiene 300 registros que representan viajes en el sistema de transporte masivo. Las variables incluidas fueron: identificador de usuario, hora del día, estación de origen, estación de destino, tiempo de viaje (en minutos) y tipo de transporte (bus, metro o tranvía). Estas variables permitieron aplicar técnicas de agrupamiento y descubrir estructuras latentes en los datos.

# Desarrollo del Modelo

Para el análisis se utilizó el algoritmo de agrupamiento K-Means. Primero, las variables categóricas fueron transformadas mediante codificación numérica con LabelEncoder. Posteriormente, se determinó el número óptimo de clusters mediante el método del codo, concluyendo que 3 clusters era una opción adecuada. Con este valor, se aplicó el algoritmo y se obtuvieron tres grupos distintos de comportamiento de los usuarios.

# Resultados

Los clusters formados revelaron los siguientes patrones:  
- Cluster 0: Usuarios que viajan principalmente en la tarde (13:40) con trayectos largos (~35 minutos), principalmente usando el metro.  
- Cluster 1: Viajes alrededor del mediodía, de duración media (~23 minutos), con uso mixto de transporte.  
- Cluster 2: Viajes cortos (~11 minutos), también en la tarde, usando principalmente buses.  
Estos patrones sugieren distintos tipos de uso del sistema de transporte y pueden ser útiles para mejorar la gestión de la demanda y optimización de rutas.

# Conclusión

El análisis mediante técnicas de agrupamiento no supervisado permitió clasificar eficazmente los distintos perfiles de uso del transporte masivo. La técnica de K-Means demostró ser útil para descubrir patrones ocultos, incluso en datos simulados. Este tipo de análisis puede tener aplicaciones prácticas en la planificación urbana y en la toma de decisiones para mejorar la eficiencia del transporte público.

# Referencias

Palma Méndez, J. T. (2008). \*Inteligencia artificial: métodos, técnicas y aplicaciones\*. McGraw-Hill España.  
Scikit-learn. (n.d.). \*Clustering: K-means\*. https://scikit-learn.org/stable/modules/clustering.html#k-means