



## **ANALISTA DE DATOS.**

UNIDAD 11

Aprendizaje automático.

## Objetivo

Afianzar el conocimiento obtenido en la clasificación.

## Temática

Clasificación.

## Descripción de la actividad

1.

- Usando Socrata, Cargue un dataset que se encuentre en el portal de datos públicos de Colombia
- Realice las tareas de limpieza necesarias
- Dividir el conjunto entre datos de entrenamiento y datos de prueba, los datos de prueba serán el 15% del dataset
- Aplicar el algoritmo de clasificación de los K vecinos más cercanos
- Calcular los puntajes de la clasificación
- Concluir si la clasificación fue buena o no

2.

- Usando Socrata, Cargue un dataset que se encuentre en el portal de datos públicos de Colombia
- Realice las tareas de limpieza necesarias
- Aplicar el algoritmo de agrupamiento de las K medias
- Calcular los puntajes de la agrupación
- Representar gráficamente los grupos
- Concluir si la agrupación fue buena o no

## Material requerido/ recursos de apoyo

- Video del módulo
- Cuenta creada en github
- Cliente de github instalado
- Editor de notebooks, se sugiere deepnote
- La documentación del módulo.