

Uso de Geometría y Topología para la Ciencia de Datos
Examen Argumentativo 02
Periodo FJ-2023
Grupo 602

INSTRUCCIONES GENERALES:

- A continuación se te presentan unas preguntas/ejercicios que te evaluarán las competencias disciplinares y procedimentales del bloque.
- Cada pregunta, procedimiento tiene una ponderación dada, y dependiendo tu respuesta puedes obtener el puntaje total o parcial.
- Este examen se realizará por equipos especificando la participación de cada integrante en un apéndice al final del archivo pdf.
- En este examen se evaluarán las competencias disciplinares:
 - Estructuración de datos (**SMA0301**).
 - Descubrimiento de información/exploración y análisis de datos (**SMA0302**).
 - Aprendizaje e IA (**SMA0401**).
 - Reconocimiento de patrones, lenguaje natural e IA (**SMA0402**).
- Al dar inicio a este examen, manifiesto conocer y comprender el Código de Ética y acepto tomar este examen final bajo estos lineamientos. Los trabajos expuestos a continuación son exclusivamente de mi autoría.

1. En esta primera parte se trabajará con el conjunto de datos `load_uschange` de la librería `sktime.datsaeries`:
 - a) (10 pts.) Suponga que la primer columna representan días de la semana. Realiza una serie de tiempo para cada una de las columnas.
 - b) (10 pts.) Determine si alguna de las variables es periódica, optimizando los parámetros necesarios.
 - c) (30 pts.) Realiza un informe de lo encontrado en un archivo latex, puedes incluir gráficas.
2. En esta segunda parte trabajaremos con el archivo `cyclistic_clean.csv`
 - a) (10 pts.) Crea una variable entera que enumere los distintos tipos de estaciones. Con esto crea un dataframe que contenga:
 - Estación de inicio
 - Estación final
 - Ride_time_m
 - Número de viajes de estación a estación
 - b) (10 pts.) Con este dataframe crea un grafo en Mapper en los siguientes casos:
 - Función de proyección en las variables Número de viajes y tiempo de viajes
 - Usar algoritmo de clusterización KMeans con 3 clusters
 - Usar una cubierta de 3 elementos con un 30 % de traslape
 - c) (30 pts.) Realiza un informe de lo encontrado en un archivo latex, puedes incluir gráficas. Obtén información sobre las distintas componentes conexas de la gráfica y su relación con cada grupo de estaciones. Dar interpretación de los datos relacionado a las conexiones entre los nodos del grafo.

Archivos entregables:

- 2 Códigos de python (uno por parte del ejercicio)
- Informe escrito en latex, formato pdf.