

Práctico 2 - Preprocesamiento

Parte 1 Preprocesamiento de datos

Seleccione uno de los siguientes datasets de OpenML y determine qué problemas presenta y qué tareas de preprocesamiento pueden ser aplicadas:

- Dataset # 17 (bridges)
- Dataset #188 (eucalyptus)

Cree un nuevo cuaderno en Google Collaboratory para estudiar el dataset elegido y realice sobre el mismo las tareas de preprocesamiento requeridas. Para esto se deberá:

1. Cargar el dataset.
2. Realizar un análisis exploratorio de los datos.
3. Aplicar al menos 3 tareas de pre-procesamiento que crea necesarias (datos faltantes, discretización de valores, one hot encoder, corrección de escala, reducción de datos, etc.).

No solo se debe incluir en el cuaderno el código, sino además una breve explicación que justifique las decisiones tomadas.

Fecha de entrega: 18/04/2021