

Actividad Acumulación de puntos

1.- Mediante el uso de Python y Jupyter Notebooks, debe desarrollar una rutina que permita ejecutar las siguientes tareas:

- Cargue los datos en un dataset de Pandas.
- Identifique que columnas deben ser tratadas.
- Proponga una rutina de limpieza de datos.

Trabaje con el Dataset: 1.4.6_Forest.csv, el cual puede encontrar en el ambiente virtual o adjunto en el email.

2.- Basándose en la rutina de limpieza y carga de datos del paso N°1 genere una nueva rutina que permita tomar el trabajo ya realizado y efectuar:

- Extracción de métricas relevantes desde los datos
- Conjunto de al menos 10 gráficos que permitan entender de mejor manera los datos.
- Hallazgos relevantes con respecto a los datos.

3.- Por último, basándonos en la rutina de exploración de datos completada en el paso N°2, determine lo siguiente para poder tener modelos de mejor calidad:

- Determine que variables son más relevantes para un posible modelo de ML
- Justifique su selección con gráficos y datos

Enviar al email: is.naranjo@profesor.duoc.cl

Día jueves 09/04 hasta las 12:00 AM