

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS E INGENIERÍA
INTELIGENCIA ARTIFICIAL INDUSTRIAL (1MTR19)
Primer Laboratorio
(Semestre 2024-0)

Instrucciones para el Laboratorio de Análisis Exploratorio de Datos

Objetivo: Realizar un análisis exploratorio de datos (EDA) exhaustivo en el conjunto de datos proporcionado del Desafío de Datos PHME 2022 (solo data: 'SPI_testing_0.csv.zip') para descubrir patrones y obtener información sobre la detección de fallas en circuitos eléctricos.

Tareas:

1. **Visión General de los Datos (5 puntos):**
 - Cargar e inspeccionar el conjunto de datos.
 - Identificar las dimensiones del conjunto (número de filas y columnas).
 - Identificar los tipos de datos (integer, object, float, etc).
 - Determinar los tipos de variables (categóricas, continuas, etc.).
 - Establecer el identificador único de la data.
2. **Limpieza de Datos (3 puntos):**
 - Identificar y tratar cualquier anomalía o valor atípico en las variables 'Volume(%)' y 'Area(%)'.
 - Verificar la existencia de valores faltantes y manejarlos adecuadamente.
3. **Análisis Estadístico (2 puntos):**
 - Calcular y mostrar en una tabla las estadísticas resumidas (media, mínimo, máximo, desviación estándar, etc.).
4. **Visualización de Datos (6 puntos):**
 - Crear tres gráficos de barras de frecuencias relativas para tres variables categóricas y tres gráficos de densidad para tres variables numéricas continuas. Seis gráficos en total. Comentar cada gráfico.
 - Crear un 'Correlation Heat Map' para examinar las correlaciones entre las variables y comentar los hallazgos.
5. **Generación de Conclusiones (4 puntos):**
 - Resumir los hallazgos clave del análisis.
 - Discutir las posibles implicaciones de estos hallazgos.

Entregables:

- Un cuaderno Jupyter (.ipynb file) que contenga todo el código, comentarios y visualizaciones.
- Un breve informe que resuma los hallazgos e insights clave de tu EDA (en formato PDF o Word).

Fecha de Entrega: 23/01/2024, 06:00 pm (la entrega es personal).

Enviar a: john.taco@pucp.edu.pe

Nota: Asegúrese de que su análisis sea claro, bien estructurado y comunicado eficazmente.

Link solo para especificaciones de la data (no descargar data de este link): [Data Challenge 2022 – PHME24 \(phm-europe.org\)](https://phm-europe.org/DataChallenge2022)