



Análisis de datos exploratorias

EDA

Librerías: Versión

- sweetviz
- pandas
- perfilado
- Predicción diferida (lazy predict)
- FLAML (llumca)
- Pycaret

Explicación:

! Dulceviz → informe se generan en formato HTML o Jupyter notebook

→ Incluye: Correlación (Heatmaps)

- Histogramas.
- Entre otras.

- También es posible personalizar el informe

! Perfil de Pandas → utilizando la librería de pandas profiling

→ Incluye: más funciones que Dulceviz.

- Descripción general.
- Sección de alertas.
- Reproducción.
- Variables.
- Interacciones.
- Correlaciones.
- Valores faltantes.
- Muestra.

! Predicción → Aplicación de diferentes modelos para clasificar.

! FLAML → Permite comparar diferentes modelos de aprendizaje automático, ayudando en selección de hiperparámetros.

→ Incluye: FLAML

- LightGBM
- XGBoost
- Random forests
- Entre otros. → También se puede agregar modelo y restricciones.

! PyCaret -> Herramienta de gestión de modelos de un extremo a otro que nos ayuda a automatizar los procesos de aprendizaje automático.

incluye:

- Análisis con normalización, escalado y selección de características mediante PCA.
- Eliminación de valores atípicos.
- Aplicación en AWS, GCP o AZURE
- Se puede guardar en un objeto pickle.