

Source Code Directory (src)

Este directorio contiene todo el código fuente del proyecto. La estructura del directorio está organizada de la siguiente manera:

```
src/
├── data/                # Directorio para datos (raw, processed, etc.).
                        En caso de archivos grandes, trabajaremos con S3/GCP.
├── notebooks/          # Jupyter notebooks para análisis y
                        experimentación
├── models/             # Modelos entrenados y código relacionado
├── utils/              # Funciones de utilidad y helpers
└── visualization/      # Código para visualización de datos
```

Estructura de Directorios

data/

- Contiene todos los conjuntos de datos utilizados en el proyecto
- Subdirectorios:
 - **raw/**: Datos sin procesar
 - **processed/**: Datos procesados
 - **interim/**: Datos intermedios

notebooks/

- Jupyter notebooks para análisis exploratorio
- Experimentos y prototipos
- Documentación de procesos

models/

- Modelos de machine learning
- Scripts de entrenamiento
- Funciones de predicción

utils/

- Funciones auxiliares
- Código compartido entre diferentes partes del proyecto
- Herramientas de preprocesamiento

visualization/

- Código para generar gráficos
- Dashboards

- Reportes interactivos

Convenciones de Código

1. Usar Python 3.8+
2. Seguir PEP 8 para estilo de código
3. Documentar funciones y clases con docstrings
4. Mantener notebooks limpios y bien documentados
5. Incluir requirements.txt o environment.yml para dependencias