

Dirección gitlab: '<https://gitlab.com/uah3135427/TFM>'

Orden lógico , respecto a la realización del trabajo, y contenido de los notebooks y archivos python utilizados.

- DOWNLOAD_DATASET.ipynb: Descargamos el dataset desde la API del Ayuntamiento
- DOWNLOAD_SELENIUM.py: Me permite, mediante Selenium, la descarga de las categorías y subcategorías de la página web de quejas del Ayto.
- TRATAMIENTO_TRAIN_RECORDSET: Procedo a la reclasificación y agrupación de los service_code del dataset
- EDA.ipynb: Creo tablas y gráficos exploratorios
- ESTUDIO_DATASET.ipynb: preprocesamiento y limpieza del dataset train
- TOPIC_MODELLING.ipynb: creo los topics.
- PREDICCION_SOBRE_DATASET.ipynb: realizo la predicción de service_code del dataset test
- Los archivos CREAR_CLAVES.py, SIN_COORDENADAS.py, CON_COORDENADAS.py: me sirven para generar un archivo de palabras únicas en el callejero, extraer direcciones y asignar coordenadas con geocoder respectivamente.
- El archivo FOLIUM.ipynb, me permite la representación geográfica del dataset sobre un mapa.

Todos los archivos fuente y generados en la realización del trabajo están almacenados en el directorio content, en un archivo rar.

Los notebooks han sido desarrollados en Google Colab, los archivos python con Microsoft Visual Studio Code a nivel local.