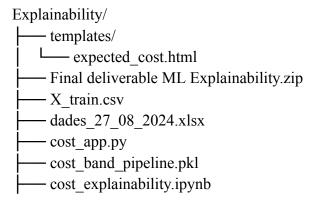
EXPLAINABILITY (SHAP AND LIME)

1. Folder structure



expected cost.html should be in a folder called templates for cost app.py to function.

2. Documents included

expected_cost.html: Interactive front-end UI in Catalan for cost prediction and SHAP/LIME explanations

Final deliverable ML Explainability.zip: packaged folder with all files for easy download

X train.csv: extracted training data used to compute SHAP/LIME

dades 27 08 2024.xlsx: initial dataset, used for model training

cost app.py: Flask backend serving UI + API for predictions and explanations

cost_band_pipeline.pkl: Pickle file containing trained sklearn pipeline with preprocessor and classifier

cost_explainability.ipynb : training of 1 model used to run and generate SHAP and LIME
(includes graphs)

3. How to run the web app

Step 1: Install requirements

pip install flask shap lime pandas numpy scikit-learn joblib

Step 2: from root Folder (either "Explainability" if downloaded from GitHub, or "Final deliverable ML Explainability" if zip folder downloaded), run the following code

python cost_app.py

! Note: it could also be python3 cost_app.py depending on the Python version that was installed

Step 3: Open the link in the terminal using a browser to visualize the frontend Typically, the URL is http://127.0.0.1:5000/

4. Backend Flask API

Endpoint: POST /api/predict Input: JSON with features key

Output: Cost band, confidence %, SHAP summary, and LIME summary

```
Example payload:
   "Exitus": 0,
   "Especialitat d'acte mèdic": 34,
   "Risc": "No",
   "Àrea assistencial": 35,
   "Tipus de praxi": "AMB CIRURGIA",
   "Centre docent": "No",
   "Àmbit": "SISCAT",
   "Consentiment informat": 0,
   "Especialitat": "OBSTETRICIA I GINECOLOGIA",
   "Centre": 11979.0,
   "Codi diagnòstic": "O813",
   "Codi procediment mèdic": 72,
   "Sequeles": 73,
   "Reclamants": 1,
   "Pacients": 1,
   "Reclamants Exitus": 0
```

5. How to use the Jupyter notebook

Step 1: ensure that the file path for the initial dataset is correct The initial dataset is dades_27_08_2024.xlsx

Step 2: run all cells to predict the cost range and show SHAP and LIME data