Explaining the workflow:

The 'train.csv' and 'test.csv' files must be run inside the 'cleaning.py' script. This .py file returns an already clean and pre-processed train, validation and test files named 'train_script.csv', val_script.csv' and 'test_script.csv' respectively. They will be saved in the 'documents' folder.

The 'train_script.csv' and 'val_script.csv' files run inside the 'search_model.py' script. Then through a for loop it will iterate over differents models in order to find the one that performs best. Penalties for 'false positives' and gains for 'true positives' will be taken into account. The model with the highest 'auc_score' and with the corresponding 'threshold' that maximizes the gain will be chosen. Finally the 'model_search.py' script will return two pkl files: 'model.pkl' which is the already trained model and the 'threshold.pkl' file. They will be saved in the 'src' folder.

Finally, the 'test_clean.csv' files along with 'model.pkl' and 'threshold.pkl' run inside the 'clients_predict.py' script. This script will return a file called 'clients_list.xlsx' with the list of clients to call to offer insurance for their vehicle.They will be saved in the 'documents' folder.

```
train.csv ──────┐                                    ┌────── test.csv
                ↓                                    ↓
              cleaning.py
         ┌────────┼──────────────────────────────────┐
         ↓        ↓                                   ↓
  train_script.csv   val_script.csv           test_script.csv
         │        │                                   │
         └────────┼──────┐                            ↓
                  ↓      ↓                     clients_predict.py
             model_search.py                         │
                  │                                   ↓
         ┌────────┤                            clients_list.xlsx
         ↓        ↓
     model.pkl  threshold.pkl
         │        │
         └────────┴──────────→ clients_predict.py
```