

Documentation technique

Dans le cadre de notre projet, nous avons créé plusieurs modules pour délimiter les différentes tâches que nous avons réalisé lors de notre projet. Cela permet de donner plus de visibilité et une meilleure compréhension de notre projet. Au sein de cette documentation, on va décrire le rôle des différents scripts utilisés dans notre projet.

1. Nettoyage des données

La toute première étape de notre projet a été de nous assurer que notre base de données est propre. Il est important d'effectuer une grande étape de nettoyage avant de se lancer dans la modélisation. Pour la partie nettoyage, elle se trouve dans le notebook [easydate.ipynb](#).

Pour la partie Data Cleaning, On a commencé par supprimer les variables que l'on a jugé peu utiles (Différentes variables identifiants et des variables en doublons.). On a ensuite nettoyé nos données quantitatives. Le format des nombres décimaux était parfois différents (Parfois « , » ou « . ») Pour terminer nous avons géré les valeurs manquantes en fonction du type de variables. On a exporté nos données nettoyées dans le fichier csv : [data_cleaned.csv](#)

Notre notebook contient également une partie Analyse de variables dans laquelle on a réalisé des statistiques descriptives.

2. Modélisation

La partie modélisation se trouve également dans le notebook [easydate.ipyn](#). Pour l'échantillonnage nous avons décidé de scinder nos données de la sorte : 70% dans le jeu de données d'apprentissage et les 30% restant dans le jeu de données test. On ensuite appliqué la méthode SMOTE sur nos données d'apprentissage.

On a décidé d'implémenter un modèle prédictif avec la méthode Randomforest. On a choisi les paramètres à tester avec GridSearch.

3. Application Dash

Nous avons créé une application afin de représenter graphiquement les données de notre base. Nous avons créé plusieurs scripts pour les différents graphiques.

[dash-bar.py](#)

[dash-piechart.py](#)

[dash-scatter.py](#)

[dash-result.py](#)

4. Déploiement sur Heroku

Le fichier Heroku-dash donne accès aux liens pour accéder à l'application sur Heroku