

Document de cadrage : Projet data MD5

Mise en contexte : Ce module a pour objectif de mettre les étudiants face à des vraies problématiques d'un projet data. Ces problématiques traitées couvriront plusieurs thématiques : Récupération de données, analyse de données, modélisation, industrialisation, mise en valeur de la solution...

Problématique :

L'objectif de ce projet est de mettre en place un outil d'estimation de véhicules d'occasions.

Plus généralement, l'estimation d'un prix de marché d'un produit est une problématique récurrente dans les cas d'usages data. Exemples : Estimation d'un prix d'un logement, estimation d'un prix d'un véhicule, etc.

Objectif :

A partir de la donnée fournie, l'objectif est de construire un modèle capable de prédire le prix de vente : le prix de marché d'un véhicule d'occasion (à la manière de l'exemple vu en cours : <https://www.aramisauto.com/reprise/estimation>). Où le modèle prend en input les caractéristiques du véhicule et sort en output le prix du véhicule.

Déroulement :

Les étudiants auront la liberté de choisir l'approche et la modélisation qui leur semblent pertinentes.

Tout au long ce projet, des séances seront dédiées à l'avancement en autonomie.

Cependant, des séances régulières seront planifiées afin d'aider les étudiants à avancer sur le projet et cerner les problématiques métiers sous-jacentes.

Aussi, des points individuels seront à prévoir afin de mieux orienter les étudiants sur les choix techniques.

NB : Les étudiants ne seront pas limités à l'utilisation des données fournies mais pourront l'enrichir avec de la donnée externe dans le but d'affiner les prédictions.

Organisation :

Pour ceux qui le souhaitent, il est possible de faire ce projet en binômes.

Les binômes devront être communiqués au plus tard le 28 avril et ne pourront plus être modifiés.

Rendu :

A la fin du projet (deadline à définir), les étudiants devront rendre :

- Le code du projet (sous la forme d'un notebook ou git repository...)
- Un rapport dans lequel ils détailleront la solution : choix techniques et métiers ainsi que les résultats obtenus. Plus de détails sur le rapport seront communiqués ultérieurement.