**Introduction**

Le projet consiste en la création d'un modèle de machine learning pour vérifier la solvabilité d’un client pour un prêt en fonction de diverses caractéristiques. Le modèle est entraîné sur un ensemble de données historiques sur la vie du client ainsi que son passif bancaire.

Le modèle est entraîné et évalué à l'aide de la bibliothèque Scikit-Learn, et les expériences sont suivies et enregistrées à l'aide des bibliothèques MLflow, joblib et pickle. Les modèles entraînés sont stockés dans un serveur centralisé, également géré par streamlit.

Le modèle est ensuite déployé sous forme d'une API qui peut être utilisée pour vérifier cette solvabilité en choisissant un client précis.

Un dashboard interactif est également créé pour visualiser les prédictions du modèle et analyser les résultats.

**Structure du projet**

Le projet est organisé en plusieurs dossiers :

* **data/** : contient les fichiers de données brutes, ainsi que les données prétraitées utilisées pour l'entraînement et l'évaluation du modèle.
* **models/** : contient les différents modèles entraînés, stockés sous forme de fichiers de modèle Scikit-Learn, ainsi que les informations sur les expériences d'entraînement et d'évaluation stockées via MLflow.
* **dashboard/** : contient le code pour le dashboard interactif.
* **api/** : contient le code pour l'API permettant de prédire les prix des maisons.

Chacun de ces dossiers contient un fichier **README.md** expliquant le contenu et la structure du dossier.

**Packages utilisés**

Les packages Python utilisés pour ce projet comprennent :

* numpy
* pandas
* scikit-learn
* mlflow
* joblib
* pickle
* xgboost, lightgbm, linearreg, randomforest, …
* flask
* plotly