

Projet Examen – Machine Learning

Titre : Immo Predictor : Valorisation et Diagnostic Intelligent

Dataset : <https://www.kaggle.com/datasets/lespin/house-prices-dataset/data>

Objectif : Développer une plateforme capable :

1. D'estimer le prix d'un bien immobilier (Régression).
2. De classer automatiquement son type (Classification).

- Partie 1

Régression

Cible : SalePrice

Variables utilisées : GrLivArea, TotalBsmtSF, LotArea, BedroomAbvGr, FullBath, TotRmsAbvGrd, OverallQual, OverallCond, YearBuilt, YearRemodAdd, Neighborhood, GarageCars, GarageArea, PoolArea, Fireplaces.

- Partie 2

Classification

Cible : BldgType

Variables utilisées : GrLivArea, TotRmsAbvGrd, OverallQual, YearBuilt, GarageCars, Neighborhood, HouseStyle.

Étapes demandées :

1. Analyse Exploratoire des Données (EDA)
2. Pré-traitement :
 - a. Gestion des valeurs manquantes Encodage des variables catégorielles
 - b. Normalisation / Standardisation si nécessaire

3. Modélisation :

- a. Régression : Decision Tree, Random Forest
- b. Classification : SVM, Random Forest
- c. Évaluation :
 - i. Régression : MAE, RMSE, R^2
 - ii. Classification : Accuracy, F1-score, Matrice de confusion

Livrable :

Les étudiants doivent envoyer :

- Un Notebook propre et structuré contenant le code.
- Un court rapport (2 à 3 pages maximum).

Le fichier compressé doit obligatoirement être nommé :
prenom_nom_classe.zip

Date limite : 02/02/2025

Envoi par mail à : amadoualwalyndiaye@gmail.com