



SOMMAIRE



- Objectifs du projet
- Datasets
- Technologies utilisées
- Main.py
- Preprocessing.py
- Modeles.py
- Ref_Texte.py
- Démonstration
- Axes d'amélioration





OBJECTIFS DU PROJET

- Récupération DB PostgreSQL
- Elaboration Application Machine Learning (Streamlit)
 - Prétraitement général des données
 - NaN Outliers Encodage Standardization
 - Split train/test
 - Implantation de plusieurs modèles
 - Classification & Régression
- Déploiement de l'application





EXPLORATION DES DONNEES

DATASETS: VIN

Id: Float64

fixed_acidity: Float64

volatile_acidity: Float64

citric_acid: Float64

residual_sugar: Float64

chlorides: Float64

free_sulfur_dioxid: Float64 total_sulfur_dioxid: Float64

density: Float64

ph: Float64

sulphates: Float64

alcohol: Float64

target: Int64



DATASETS: DIABETE_INDE

Id: Int64

nombre_grossesse: Int64

concentration_glucose: Int64

pression_arterielle: Int64

epaisseur_pli_cutane: Int64

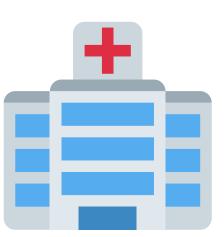
insuline_serique: Int64

imc: Float64

fct_genealogique_diabete: Float64

age: Int64

target : Objet



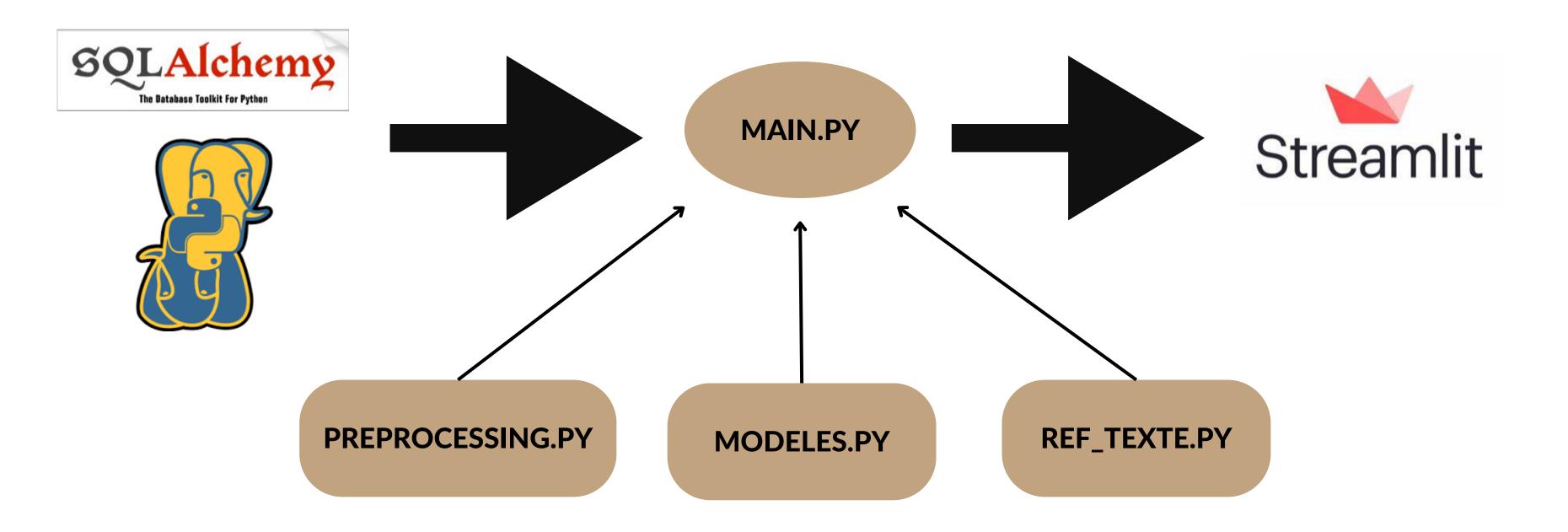


TECHNOLOGIES UTILISÉES



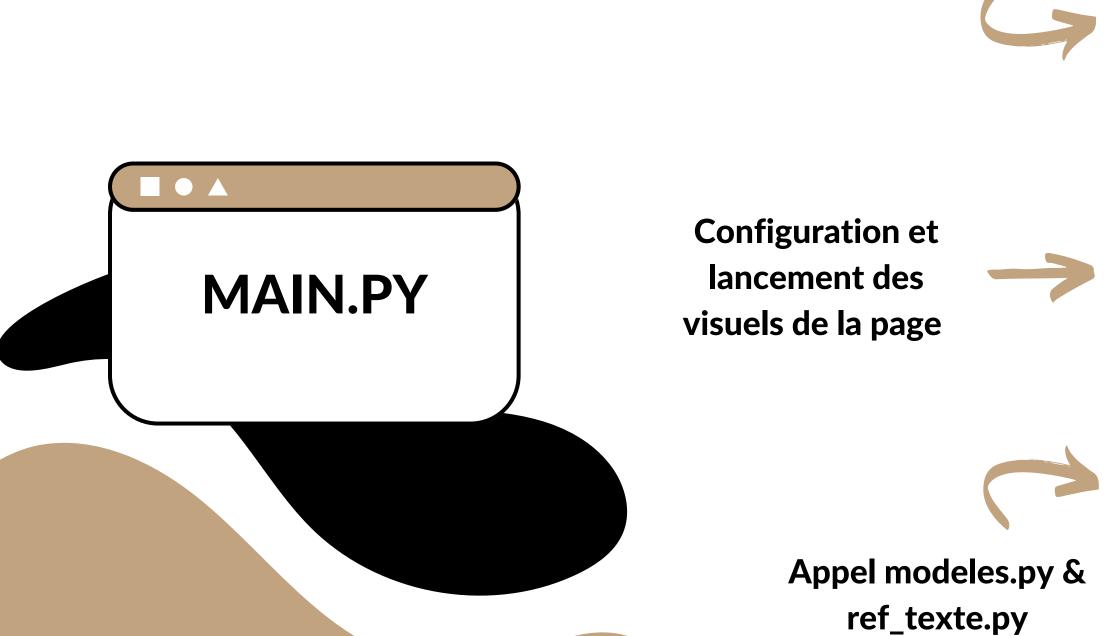


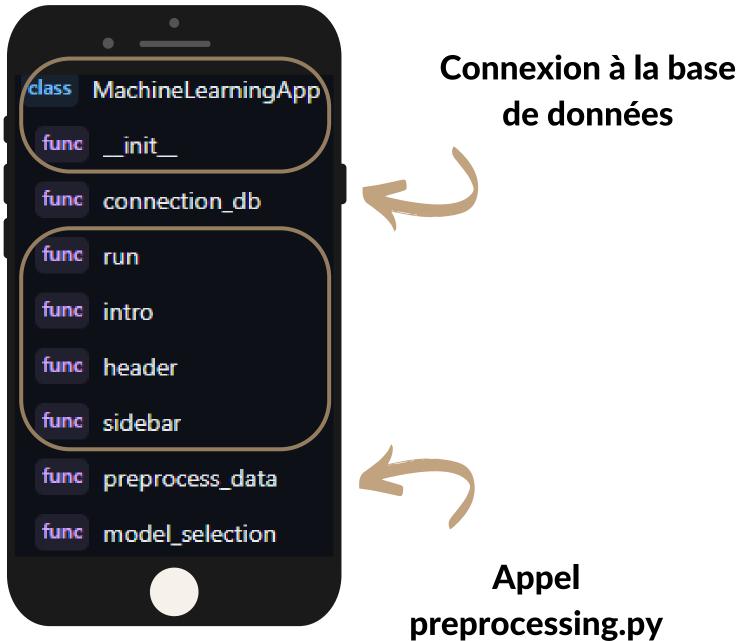
FONCTIONNEMENT CODE





Gestion et initialisation de l'application

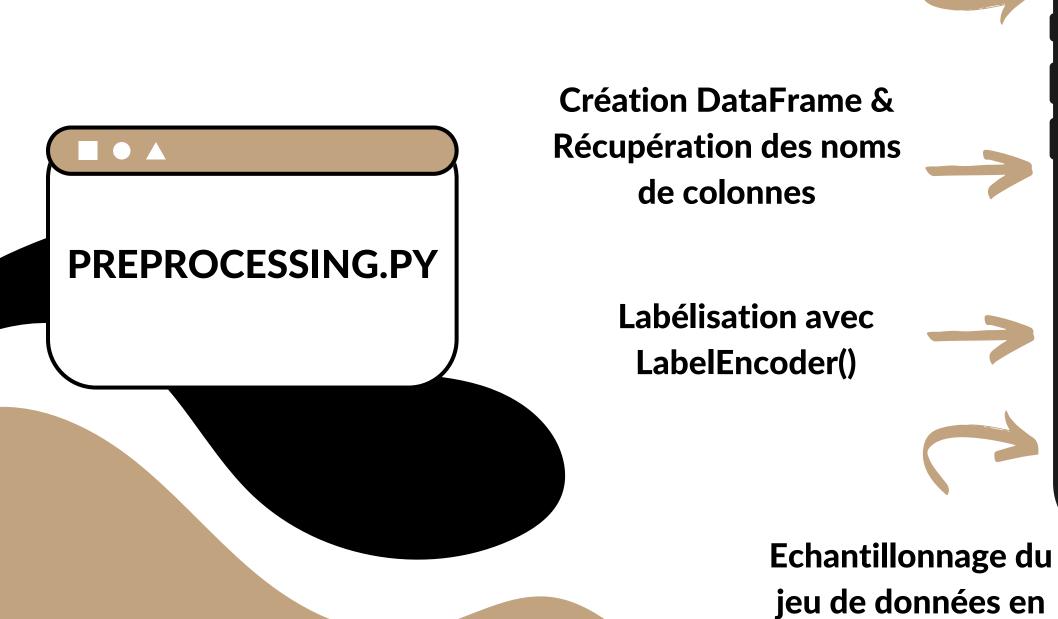


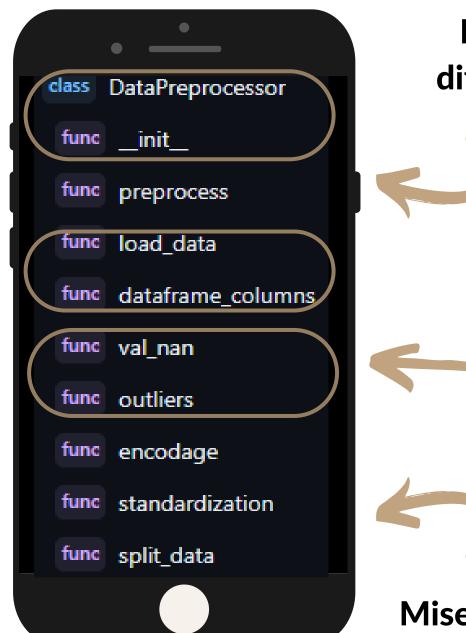






train/test





Lancement des différentes étapes

Gestion des valeurs manquantes & outliers

Mise à l'échelle des données avec StandardScaler()



MODELES.PY

Gestion des dictionnaires des modèles de ML



Gestion des modélisations et des graphiques

Sortie des métriques et des graphiques



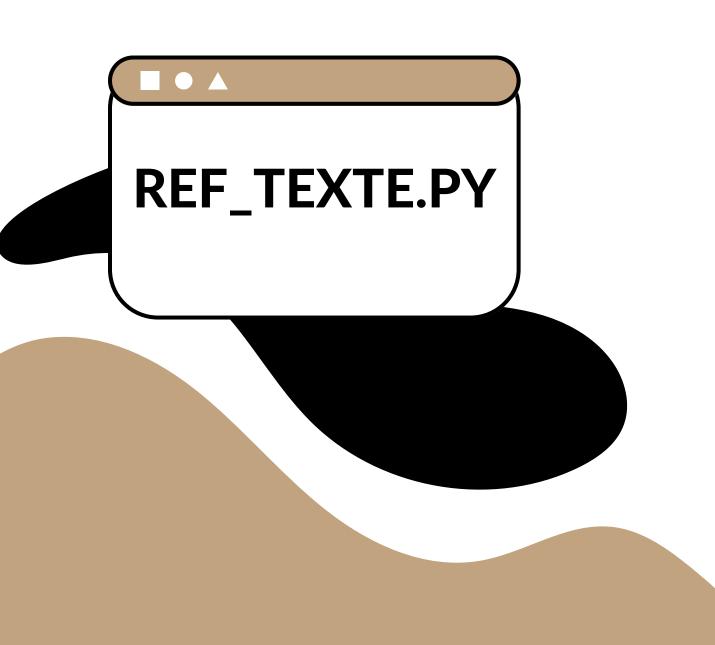
Répartition entre Régression et Classification





Modélisations des classifications ou régressions





Gestion des dictionnaires explications





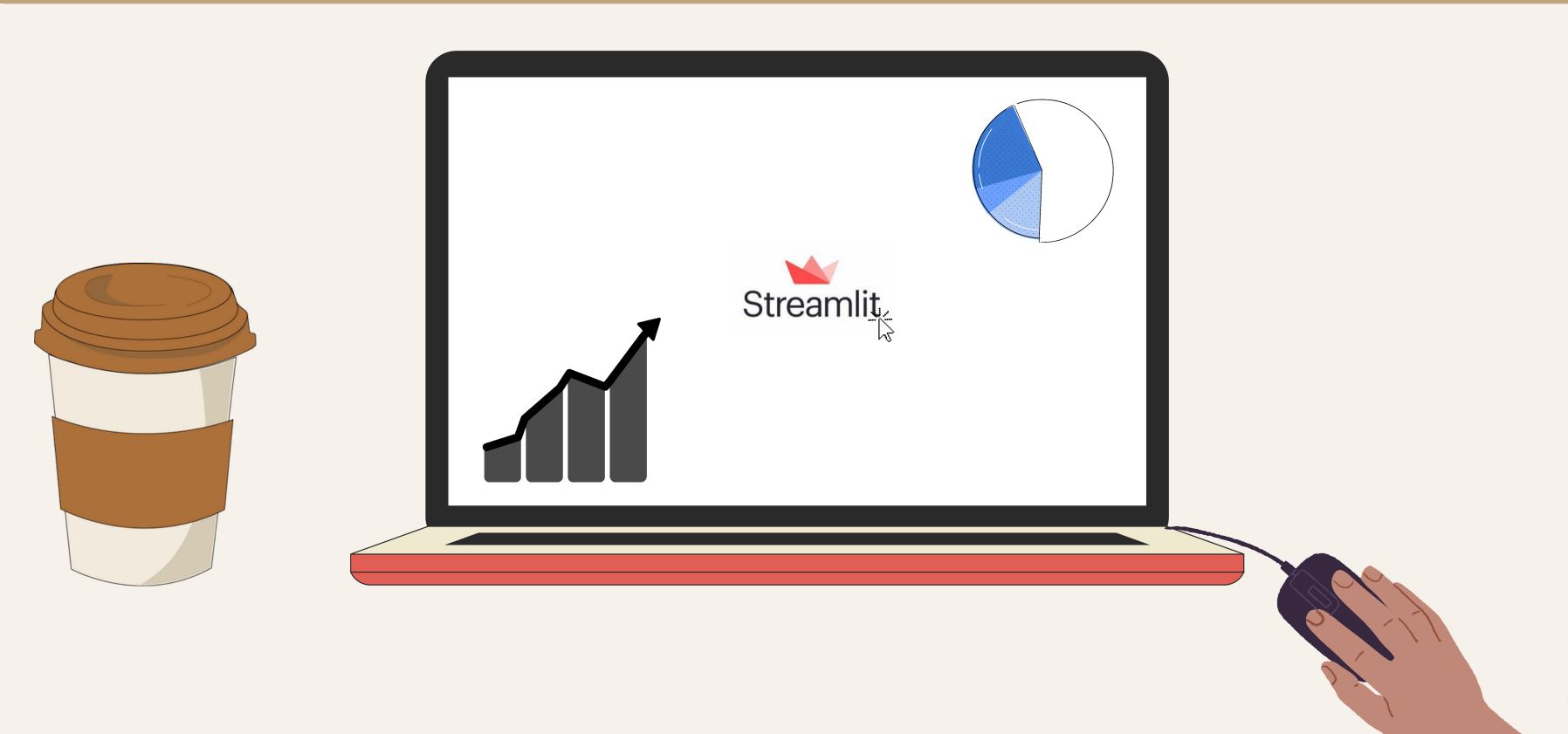
Gestion des explications des metrics



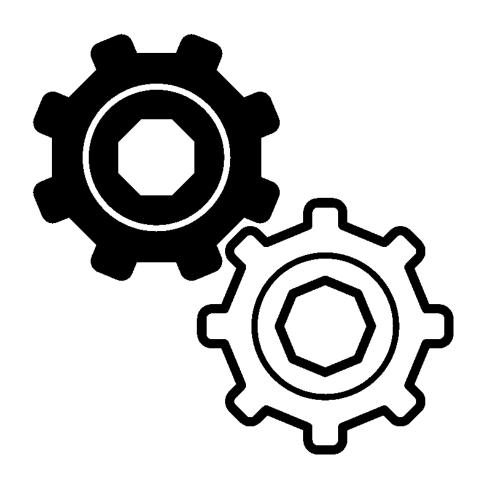
Gestion des avantages/inconvénients des algorithmes



DÉMONSTRATION







AXES D'AMÉLIORATION

- Gestion des conflits des Hyperparamètres
- Importation fichier (csv, JSON, etc...)
- Ajout du choix seuil NaN par colonnes/lignes
 - Choix gestion des outliers
- Mise en place OrdinalEncoder()
- Choix entre Standardisation et Normalisation
- Enrichissement des dictionnaires de modèles



