

Octave Pouillot

Projets de Formation

Projets de formation

- Projet 1 - Analysez des données de systèmes éducatifs
- Projet 2 - Concevoir une application au service de la santé
- Projet 3 - Anticipez les besoins en consommation de bâtiments
- Projet 4 - Segmenter des clients d'un site e-commerce
- Projet 5 - Classification automatique des biens de consommation
- Projet 6 - Implémenter un modèle de scoring
- Projet 7 - Déployer un modèle dans le cloud

Projet 1

Analyser des données de systèmes éducatifs

Contexte :



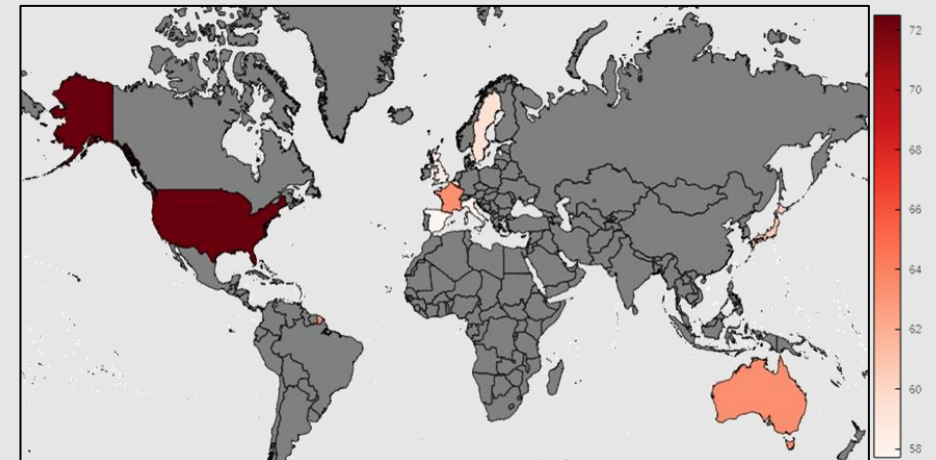
Entreprise de formation en ligne

Mission :



Exploratory Data Analysis en vue d'une expansion à l'internationale

- ✓ Sélection des indicateurs pertinents
- ✓ Etablissement d'un score par pays basé sur ces indicateurs
- ✓ Analyse prospective de l'évolution du potentiel de clients
- ✓ Conseils de stratégie basé sur les scores, avec regard critique



Projet 2

Concevoir une application au service de la santé

Contexte :



Appel à projet : Etude de faisabilité d'une application en lien avec l'alimentation

Idée d'application :



Application de suivi des apports nutritionnels pour sportif végétarien

- ✓ Sélection des indicateurs pertinents
- ✓ Imputation k-NN (nutriscore)
- ✓ Analyses bivariées (matrice de corrélation, ANOVA, Chi-2)
- ✓ Analyses multivariées (dendrogramme, ACP)
- ✓ Conclusion critique sur la faisabilité et les limites du projet



Projet 3

Anticiper les besoins en consommation de bâtiments

Contexte :



Ville de Seattle

Mission :



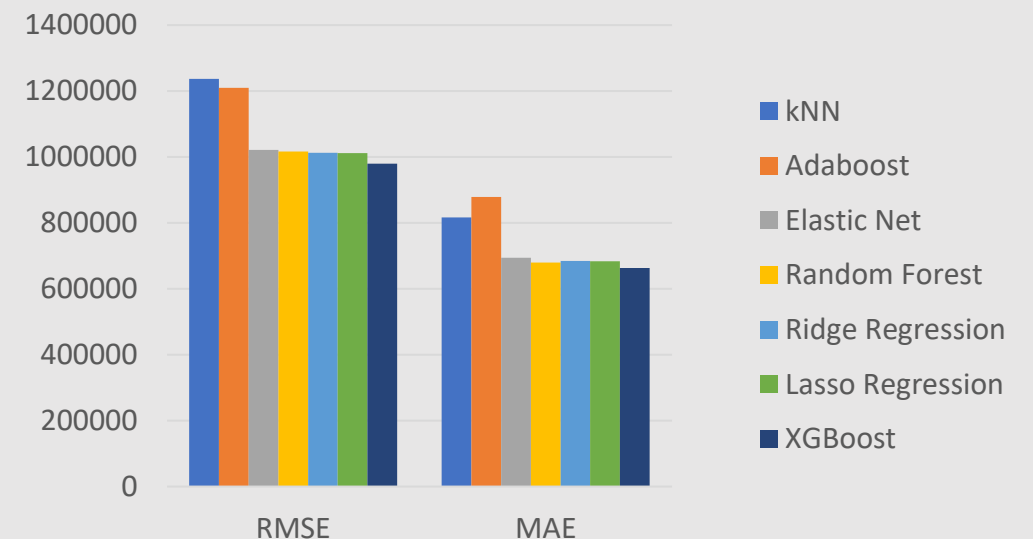
Prédire la consommation énergétique
& émissions CO2 de bâtiments



XGBoost

- ✓ Analyse exploratoire des données
- ✓ Feature pre-processing
- ✓ GridSearch (Cross Validation, Pipeline, Hyper-parameter tuning)
- ✓ Features Importance

Energy Use – Comparaison RMSE & MAE



Projet 4

Segmenter des clients d'un site e-commerce

Contexte :

olist

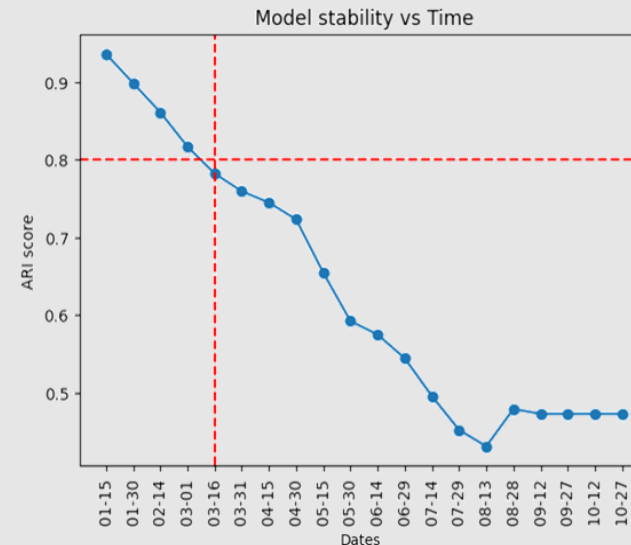
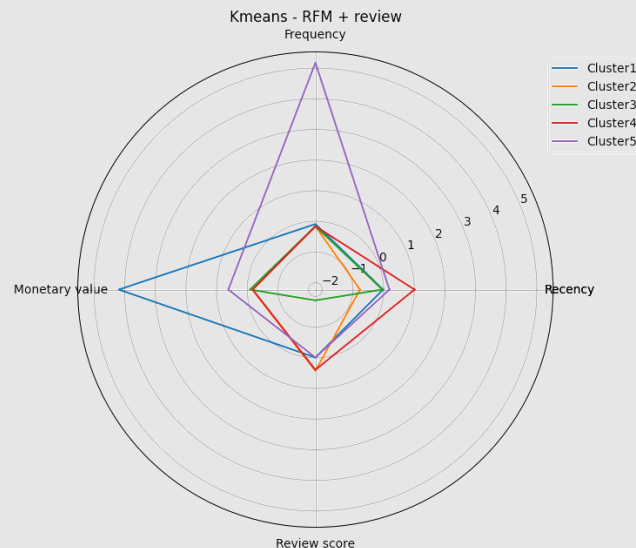
MarketPlace e-commerce

Mission :



Segmentation clients pour
l'équipe marketing

- ✓ Analyse exploratoire des données
- ✓ Featuring RFM (Récence, Fréquence, Montant)
- ✓ Clustering (kmeans, CAH, DBSCAN)
- ✓ Evaluation basé sur des scores spécifiques (Silhouette, Calinski Harabasz, Davis Bouldin)
- ✓ Simulation de DataDrift pour maintenance



Projet 5

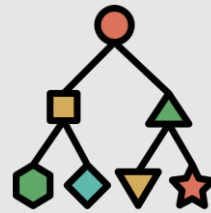
Classification automatique des biens de consommation

Contexte :



MarketPlace e-commerce

Mission :

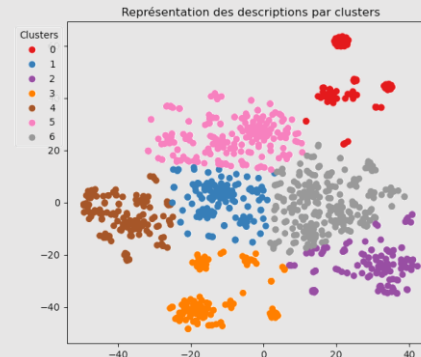
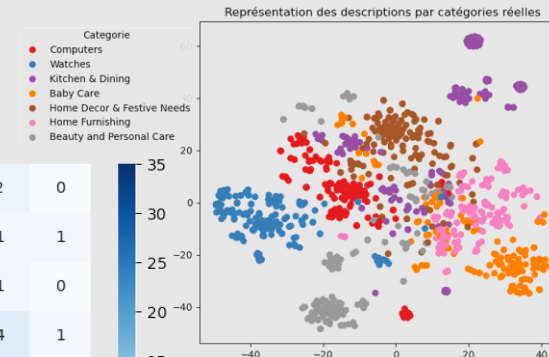


Classification automatique
d'articles via image et description

- ✓ Etude de faisabilité (T-SNE, Kmeans, ARI)
 - Par description : CountVectorizer, TF-IDF, Word2Vec, BERT, USE
 - Par photos : SIFT, CNN
- ✓ Classification supervisée d'images (CNN transfert learning)
- ✓ Test de récupération de données sur une API



Home Furnishing	26	1	0	5	4	2	0
Baby Care	2	28	4	2	0	1	1
Watches	1	1	33	1	1	1	0
Home Decor & Festive Needs	1	0	1	29	2	4	1
Kitchen & Dining	6	1	0	2	29	0	0
Beauty and Personal Care	1	0	3	2	0	32	0
Computers	0	1	1	1	0	0	35
	0	1	2	3	4	5	6



Projet 6

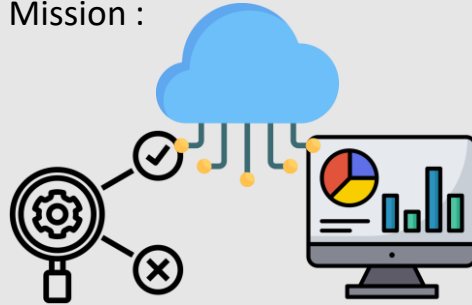
Implémenter un modèle de scoring

Contexte :



Société financière de crédit à la consommation

Mission :



Prédire la probabilité de faillite d'un client
Dashboard interactif pour l'équipe relation client
Mise en production Modèle & Dashboard

- ✓ Traitement du déséquilibre des données (SMOTE)
- ✓ Etablissement de scores métier (matrice de confusion)
- ✓ Conception d'un modèle de scoring pour prédire la probabilité de faillite d'un client (MLFlow tracking & Models)
- ✓ Réalisation d'une API (FastAPI)
- ✓ Réalisation d'un Dashboard interactif appelant l'API (Streamlit)
- ✓ Mise en production de l'api et du dashboard (CI/CD, GitHub Actions, PyTest, Azure Web App, Streamlit hosting)
- ✓ Estimation du Data Drift (EvidentlyAI)
- ✓ Documentation de la solution



Projet 7

Déployer un modèle dans le cloud

Contexte :



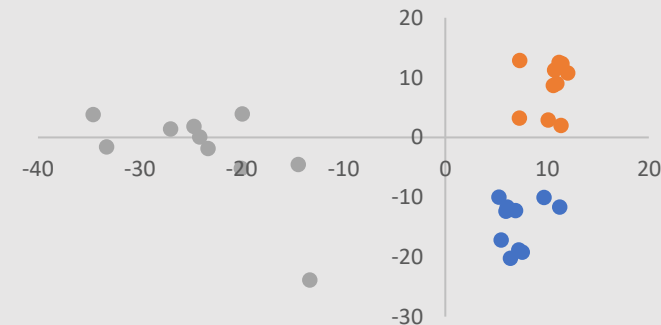
Start-up de l'AgriTech
Application mobile de reconnaissance de
fruits

Mission :



Mettre en place un moteur de classification de fruits
Compléter une chaîne de traitement
Mettre en œuvre une architecture Big Data

- ✓ Mise en place d'une machine virtuelle
- ✓ Création d'une architecture AWS
 - IAM
 - S3
 - EMR
- ✓ Développement d'une PCA PySpark
- ✓ Validation de la Proof of Concept



● Citron ● Pomme ● Banane