

- **Projet Walmart sales : optimisation des processus marketing de qualification de prospect par le biais d'algorithmes d'apprentissage supervisés (régression).**
Détermination de prix.
K-Means, Dbscan, Plotly
- **Projet Conversion rate : optimisation d'algorithmes d'apprentissage automatique supervisé (classification) sur des bases de données déséquilibrées.**
Recommandation à l'équipe produit/marketing pour l'amélioration du taux de conversion.
K-Means, Dbscan, Plotly
- **Projet UBER Pickups : localisation de zones de densité géographique par l'élaboration d'algorithmes d'apprentissage automatique non-supervisé.**
meilleurs emplacements pour un chauffeur UBER pour récupérer les clients le plus rapidement.
K-Means, Dbscan, Plotly
- **Projet AT&T Spam Detector : analyse prédictive de données non-structurées par l'intelligence artificielle**
prédiction de spams, Analyse de sentiment, par l'élaboration d'un algorithme permettant de déterminer le sentiment d'un utilisateur à l'égard d'un produit.
Pandas, Numpy, Scikit Learn, réseaux de neurones, transfert learning
- **Projet Getaround Analysis : industrialisation d'un algorithme d'apprentissage automatique et automatisation des processus de décision**
création de KPIs, d'un web dashboard, construction et mise en production d'une application web d'intelligence artificielle via un API.
Pandas, Numpy, Scikit Learn, Streamlit, FastAPI, MLFlow, Docker
- **Projet Kayak : construction et alimentation d'une infrastructure de gestion de donnée**
Construction d'une infrastructure Cloud accueillant des données Big Data (collecte de données web, intégration des données dans un Data Lake AWS S3, nettoyage et chargement des données dans une base de données type AWS Redshift par traitement parallélisé si nécessaire via la construction d'un processus ETL).
Python, Pandas, Numpy, Scrapy, API, S3, data analysis, ETL, SQLAlchemy, PGMyAdmin
- **Projet Speed dating challenge : analyse exploratoire, descriptive et inférentielle de données**
Gestion de valeurs manquantes et aberrantes d'une base de données non-massives puis analyse pour déterminer et présenter des tendances par le biais de graphiques.
Analyse descriptive des facteurs qui influencent le fait d'obtenir un second rendez-vous lors du speed dating
Python, Pandas, Numpy, Matplotlib, Seaborn
- **Projet Prévision du cours des crypto-monnaie**
Python, Pandas, Numpy, LSTM