

EXPÉRIENCES PROFESSIONNELLES

MOE & Data Scientist

08/2021 – Aujourd'hui

CAB FORMATIONS

Tremblay-en-France

- Recueillir et analyser les besoins
- Automatiser les processus manuels avec Python et VBA
- Développer un ETL pour la prospection B2B (web scraping, API, JSON)
- Mettre en place les workflows data sous le Cloud de Zoho
- Enrichir le Datawarehouse dans le cadre de projets data driven
- Développer et optimiser les requêtes SQL
- Construire les dashboards BI (Dataviz, KPI, reporting)
- Proposer et mettre en place des modèles de prévisions robustes
- Assurer la formation des utilisateurs finaux

Outils: Python, Pandas, BeautifulSoup, Pyspark, SQL, BigQuery (GCP), Docker, Zoho, Power BI, Excel, Asana (gestion de projets)

Ingénieur Data Scientist

11/2020 – 04/2021

ANFR

Maisons-Alfort

- Recueillir l'expression des besoins et participer à la rédaction du cahier des charges
- Développer un outil pour envoyer des alertes automatiques basées sur les données
- Modéliser, concevoir les bases de données, créer les tables
- Réaliser le processus d'alimentation des bases de données avec l'outil ETL Pentaho (transformation, restitution, ingestion des données)
- Réaliser les tests et faire le choix des algorithmes de machine learning sous R avec le package Caret. Comparer les modèles de prédiction
- Créer les applications web pour déployer les modèles de prédiction
- Rédiger les spécifications fonctionnelles et techniques

Outils: Pentaho, R, Rstudio, R Shiny, dplyr, purrr, Caret, R Markdown, SVN, vi, putty, python (POO), PostgreSQL, MySQL, CentOS

Data Scientist

04/2019 – 09/2019

ANFR

Maisons-Alfort

- Identifier les sources de données pertinentes pour la modélisation statistique
 - Créer un outil pour collecter automatiquement les données météorologiques
 - Réaliser le traitement et l'analyse des données collectées
 - Assurer la disponibilité et la qualité des données
 - Prédire la perturbation de la réception de la TNT en France avec les données météorologiques (séries temporelles: ARIMA et LSTM)
 - Créer les tableaux de bord pour visualiser les résultats de la prédiction
- Outils: R, RStudio, dplyr, ggplot2, plotly, R Shiny, forecast, FactomineR, MySQL, VBA/Excel (Macros, Tableaux Croisés Dynamiques, RechercheV)

FORMATIONS

Ingénieur en Mathématiques Appliquées

Option: Statistique/Data Science
École d'Ingénieurs, Sup Galilée

2020

Université Sorbonne
Paris Nord

Master 1 Mathématiques Fondamentales

2016

Université Paris-Est Créteil

Double Licence Mathématiques-Physique

2015

Université Paris-Est Créteil

COMPÉTENCES

Machine learning/Deep learning

- Algorithmes: Random Forest, SVM, PCA, Decision trees, k-means, XGBoost, CNN, Anomaly Detection, Logistic Regression

- Librairies: Caret, ggplot2, dplyr, pandas, Numpy, Scikit-learn, Pycaret, Matplotlib, Streamlit, MLlib (Pyspark), Scrapy, Plotly, FastAPI, BeautifulSoup, Selenium, Prophet

Langages de programmation

- R, Python, PySpark, VBA, DAX, Shell/Bash

Bases de données

- MySQL, PostgreSQL, NoSQL

Systèmes d'exploitation

- Linux, macOS, Windows

Logiciels connecteurs ETL

- Pentaho, Talend, BigQuery, Supermetrics

Data visualisation & BI

- R Shiny, Power BI, Python Streamlit, Tableau, Superset, Zoho Analytics, Looker Studio

Environnements Big Data & Cloud

- Google Cloud Platform, AWS, Spark, Kafka

ATOUTS

Dynamique

Rigoureux

Curieux

Esprit d'équipe

Motivé

Autonome

LANGUE

Anglais (TOEIC 785)

