

Expériences Professionnelles

GIFI – Tech Lead ML

Sur site, **France**
Septembre 2023 – En cours

- ❖ Mise en Place d'une Plateforme Data Science de Bout en Bout
 - ❖ Développement d'un Pipeline de Données en Temps Réel
 - ❖ Déployer un Projet d'IA Générative pour l'Automatisation des KPIs & Reporting
 - ❖ Participé activement à l'organisation de l'équipe : Daily meetings, tâches de management, et recrutement.
- Environnement technique** : AWS, GitLab, Jira, Python, CI/CD.

Banque de France Freelance : Tech Lead – Data Scientist

Hybride, **France**
Février 2022 – Avril 2023

Machine Learning for security

- ❖ Conduit les revues de backlog Machine Learning pendant les phases de construction (Build) et de livraison (Delivery) d'un système en temps réel. Le système attribue les incidents aux groupes appropriés pour résolution, et met à jour ces informations toutes les 5 minutes sur un tableau de bord Splunk.
- ❖ Audité un modèle de Machine Learning attribuant un score de priorité aux incidents pour une gestion plus efficace des incidents en temps réel.
- ❖ Encadré deux alternants et participé activement à la gestion quotidienne de l'équipe.

Environnement technique: Splunk, Splunk Machine Learning Toolkit, Kanban, Jira, GitLab, NLP, TF-IDF, Python.

Business & Decision – Cabinet de conseil en data Consultant Data Scientist – Tech Lead

La Défense, **France**
Mars 2018 – Janvier 2022

Tech Lead - Data Scientist for Security chez Orange

- ❖ Dirigé la phase de construction (Build) et de livraison (Delivery) d'un système de classification de paquets de réseau, en utilisant des techniques de Deep Learning avancées comme les auto-encodeurs.
- ❖ Conçu et mis en place une architecture Big Data pour la collecte et la clusterisation en temps réel des paquets de réseau.

Environnement technique : Python (pandas, sklearn, Nltk, Gensim, Keras, Tensorflow), Git, Wireshark, CICFlowMeter, Power Point, Protocol diameter, Kafka, Spark.

Data Scientist chez Orange

- **Développé un outil de prévision du chiffre d'affaires (CA) au niveau marché**

- ❖ Élaboré un modèle de Machine Learning pour fournir des prévisions mensuelles du chiffre d'affaires sur une période semestrielle et annuelle, améliorant ainsi la précision des prévisions.
- ❖ Collaboré avec le métier pour consolider les prévisions au niveau de l'agence et développé une solution industrialisable basée sur le Machine Learning pour une prédiction plus précise du chiffre d'affaires.

Environnement technique : Python (pandas, sklearn, XGBoost, LSTM), Git, Excel, Power Point.

Data Scientist chez Edenred

- **Employé le Machine Learning pour améliorer la rétention client et la prospection**

- ❖ Piloté l'usage de techniques de Machine Learning pour optimiser la rétention client et la prospection, notamment par l'audit et la fiabilisation des données et la migration des données de SAS à PostgreSQL.
- ❖ Développé un outil d'assistance à la rétention client grâce à une étude approfondie de la contactabilité, présentant efficacement les résultats aux équipes métiers.

Environnement technique : SAS, PostgreSQL, Python, Power Point, Excel, méthode agile.

Data Scientist chez Nokia

- **Implémenté une solution basée sur le Machine Learning pour automatiser les réponses aux appels d'offres.**
- ❖ Créé une plateforme web utilisant le Machine Learning pour automatiser et standardiser les réponses aux appels d'offres, renforçant la productivité des ingénieurs et assurant l'homogénéité de la qualité des réponses.
- ❖ Travaillé en collaboration avec une équipe interdisciplinaire composée de data scientist seniors, de chercheurs de Bell Labs Nokia et d'une équipe d'industrialisation en Inde.

Environnement technique : Jupyter Notebook, NLP, S3, Java.

- **Optimisation du temps de résolution des tickets d'incidents techniques relevés par les clients.**
- ❖ Déployé un pipeline de traitement du langage naturel (NLP) pour optimiser le temps de résolution des tickets d'incidents techniques.
- ❖ Amélioré la solution existante grâce à l'entraînement de modèles d'embedding de mots (Word2Vec, FastText) et l'utilisation de documents techniques.

Environnement technique : Python (pandas, sklearn, nltk, gensim), Pyspark, AWS, Git.

Interne

- **Participé à la rédaction de l'offre technique/cadrage projet/meetup.**
- **Benchmarking des solutions open source et payantes pour créer un Chat-bot ressources humaines (RH)**
- ❖ Participé à la conception initiale d'un chatbot interne pour résoudre les problématiques RH.
- ❖ Comparé les performances des outils de "classification d'intentions" (RASA NLU, Recast.AI de SAP et Dialogflow de Google) pour orienter la décision sur le choix du meilleur outil pour le chatbot.
- **Environnement technique** : Python (pandas, sklearn, Nltk, gensim), RASA NLU, Recast.AI, Dialogflow, Power Point, Excel, méthode Agile.

Divers: Data Scientist - Data Analyst

VEOLIA – Data Scientist (Paris, Avril – Septembre 2017)

- **Amélioration du dispositif de tri télé opéré de déchets grâce à la classification des images hyper-spectrales**
- ❖ Réalisé un audit des trois caméras (RGB, caméra de profondeur et hyper-spectrale) et mis en œuvre des méthodes de Machine Learning pour la classification des images dans un projet multidisciplinaire.
- ❖ Mené une étude bibliographique des méthodes récentes de prétraitement et de classification des images hyper-spectrales, tout en animant des réunions et en facilitant la communication au sein de l'équipe de projet.

IM2NP, Faculté de médecine – Data Scientist (Marseille, Septembre 2016 – Mars 2017)

- **Modélisation personnalisée d'un signal électrocardiogramme (ECG)**
- ❖ Collaboré avec un cardiologue et un chercheur pour recueillir et structurer les spécifications du projet.
- ❖ Développé un modèle paramétrique de la dynamique du rythme cardiaque et conçu une interface graphique pour sa visualisation et manipulation (Matlab).

CEA Cadarache – Data Analyst (Marseille-Cadarache, Mai – Août 2016)

- **Étude du comportement multi-fractal de la ressource solaire : analyse et synthèse de données d'éclairement.**
- ❖ Lancé une étude pour analyser le comportement fractal ou multi-fractal de la ressource solaire, en utilisant des données d'éclairement provenant de panneaux photovoltaïques.
- ❖ Assuré la liaison entre deux chercheurs du centre mathématique et un ingénieur du CEA pour coordonner l'étude.

Formation

Centre de Mathématiques et Informatique – Marseille

- **Major** : Ingénierie Mathématique et Modélisation
- **Minor**: Data Science

Marseille, France

2015 – 2017

- **Mention bien, classé 2ème de la promotion.**

SKILLS

Langues: Anglais, Français, Arabe

Computer skills: AWS, Splunk, Kanban, Jira, GitLab, Python, R, SAS, Java, SQL, Amazon S3, Kafka, Spark, Git, Wireshark, Excel, PowerPoint, Word, Latex, Windows, MacOS, Linux.

Centres d'intérêt: Cyclisme, échecs.