



CONTACT

jaouenadel@protonmail.com

+33 6 20 51 96 71

FORMATION

Master Mathématiques, Actuariat, Data Sciences – Université du Maine, Institut de recherches en actuariat (IRA), LE MANS – Mention Bien
Sept. 2016 – Août 2019

Licence de mathématique appliquée aux sciences sociales (MIAHS) – Université de Rennes 2, RENNES
Sept. 2014 – Août 2015

COMPÉTENCES

- Expertise en Python
- Stack Data Science Python (pandas, numpy, scipy, scikit-learn, PyTorch, tensorflow)
- NLP, LLMs, (APIs LLMs, fine-tuning, RAG, recherche sémantique, Bert)
- SQL, noSQL
- Git
- CI/CD GitLab, docker, ansible
- Django et django-rest-framework
- Notions de développement frontend (html, css, javascript, VueJS)
- Compréhension du besoin métier
- Restitution de résultats, vulgarisation, formation
- Explicabilité des modèles de Machine learning (shap)
- Travail en équipes avec des profils variés

LANGUES

Anglais B2

Adel JAOUEN

Data Scientist freelance



Linkedin



Github

Expert Data Science Python avec solide maîtrise du cycle de vie des projets data. Expérience en conception, modélisation et implémentation d'algorithmes de machine learning. Restitution et vulgarisation des résultats. Stack data Python (numpy, pandas, scikit-learn). Spécialisation aux techniques de traitement du langage et LLMs.

Compétences en ingénierie ML et Data (Bases de données, SQL, noSQL, orienté vecteurs, ETLs, MLOps). Développement Backend (Django), frontend (VueJS) et outils de déploiement (CI/CD GitLab, Docker, docker-compose, Ansible).

Curieux et motivé, je vous propose mes compétences pour intervenir sur des projets NLP, mise en place de RAG, intégration de LLM, de la conception à la mise en production d'une solution fonctionnelle. N'hésitez pas à me contacter.

EXPÉRIENCES PROFESIONNELLES

Data scientist

Cross Data

ANGERS

Janv. 2020 – Juil. 2024

- **Conception, modélisation et implémentation** d'algorithmes de Machine Learning (NLP, LLM, RAG, recherche sémantique, vision, clustering) pour divers secteurs (industrie, énergie, assurance, agricole, etc.).
- **Création et optimisation** de pipelines de récupération et de traitement de données en production (données structurées et non structurées)
- **Mise en production** de modèles sous forme d'API REST (Django REST Framework), avec documentation Swagger. Utilisation de containers Docker et intégration dans des pipelines CI/CD (GitLab).
- **Monitoring des applications** en production avec Grafana et Métabase.
- **Requêtage d'API** pour l'usage de LLMs, ou pour enrichir les bases de données avec des données ouvertes (data.gouv.fr et INPI).
- **Développement front-end** avec Django VueJS et Plotly-Dash
- **Collaboration inter-équipes** en méthode agile sur des projets full-stack.

Alternant Data Scientist

Lab Covéa

LE MANS

Sept. 2018 – Sept. 2019

Travaux sur des données réelles de conduites récoltées à partir d'une application mobile

Stagiaire DataScience

Technicolor

RENNES

Avril 2018 – Août 2018

Travaux de recherche sur les frontières de décisions des réseaux de neurones

CENTRES D'INTÉRÊT

Musique (piano jazz au conservatoire régional d'Angers)

Sport (natation, course à pied)

Projets par secteurs d'activités



Marketing et Réputation en Ligne

- Automatisation de la réponse aux avis en ligne (intégration LLMs) sur une application web

Stack technique : Python, django, django-rest-framework, docker, docker-compose, postgresql, Vuejs CI/CD GitLab, Groq, OpenAI, Azure ML, Mistral, langchain, ollama, llama-cpp, pypdf, unstructured, huggingface, transformers, VectorDB



Audit et Réglementation

- Construction d'un RAG pour l'audit de normes réglementaires

Stack technique : Python, CI/CD GitLab, docker, docker-compose, postgresql, LLM, OpenAI, Mistral langchain, ollama, llama-cpp, pypdf, unstructured, VectorDB



Assurance et Mutuelle

- Amélioration d'un moteur de recherche ElasticSearch en intégrant des modèles de langage (Bert puis LLM)

Stack technique : Python, pandas, transformers huggingface, docker, OpenAI, fine-tuning du modèle Bert.



Énergie

Implémentation de modèles de prédiction des consommations (clustering, modèles linéaires) pour une application visant à réduire la consommation d'énergie. Mise en place de pipeline de récupération, transformation et stockage de données semi-structurées.

Stack technique : Python, pandas, scikit-learn, scipy, numpy, pandas, django, postgresql, django-rest-framework, docker, docker-compose, Vuejs, CI/CD gitlab

- Formation des équipes métiers

Stack technique : Power BI

- Formation à la migration de R vers Python

Stack technique : Python, pandas, dash, plotly, scikit-learn,



Santé animale

- Conception et mise en production de modèles de prédictions pour l'aide au diagnostic vétérinaire.

Stack technique : Python, pandas, django, django-rest-framework, MLFlow, CI/CD GitLab, docker, docker-compose, postgresql, pyodide (exécution de Python dans le navigateur)



Formation et Éducation

- Audit sur un algorithme d'aide à l'orientation professionnelle pour des élèves de terminale, en structurant le code et en améliorant sa maintenabilité.

Stack technique : Python, pandas, scikit-learn



Environnement et Déchets

- Etude statistique sur la composition de lots de déchets

Stack technique : Python, pandas, scikit-learn, numpy, scipy, powerBI



Immobilier

- Développement d'une application web pour l'estimation du prix des biens immobiliers en France

Stack technique : Python, pandas, scikit-learn, Docker, AWS cloud)

- Conception (POC) d'un modèle de contrôle du chauffage par dalle active pour un bâtiment en construction sur des données générées)

Stack technique : Python, pandas, scikit-learn, numpy, scipy



Logistique et retail

- Amélioration de la plateforme logistique en optimisant la phase de colisage (optimisation par contraintes). Conteneurisation des algorithmes et déploiement en continu

Stack technique : Docker, CI/CD GitLab, plotly-dash

- Développement d'un système de comptage automatique d'objets sur un tapis roulant (computer vision)

Stack technique : Python, pandas, opencv, pytorch, tensorflow, docker, docker-compose



Agriculture et Sylviculture

- Analyse statistique de l'impact des pratiques sylvicoles sur la qualité du bois

Stack technique : Python, pandas, plotly, scipy, numpy, scikit-learn



Restauration

- Développement d'une application de prédiction de ventes et de charges –prévisions (forecast) et séries temporelles (timeseries)– pour 750 restaurants afin d'optimiser les plannings.

Stack technique : Python, pandas, scikit-learn, numpy, pandas, docker, jenkins