

Mekbal Akram

Recherche de stage de fin d'études en Data

(69000 Lyon, France | Téléphone : +33 7 57 63 16 18 | Email : akrammekbal10@gmail.com)



Étudiant en M2 Data Science, passionné par le Data Engineering et Machine Learning, je suis en quête d'un stage de fin d'études à partir de Mars 2025. Je suis déterminé à mettre en pratique mes compétences et à contribuer à des projets innovants.

FORMATIONS

Université Claude Bernard Lyon 1 – UCBL

- Master 2 Data Science.
- Master 1 Informatique.

2024 – présent
2023 – 2024

Université des Sciences et Technologies HB – USTHB

- Master 1 Informatique Visuelle.
- Licence en Math et Informatique

2022 – 2023
2019 – 2022

EXPÉRIENCES

Administrateur système et Data management | Stage | SAIDA Automotive (4 Mois) Mars 2023 - Juin 2023

- Collaboration avec les équipes métier (ventes, logistique, stocks) pour identifier leurs besoins en données et leurs défis.
- Traduction de ces besoins en **spécifications techniques** pour les **pipelines de données** et les **tableaux de bord**.
- Conception et mise en œuvre d'architectures de traitement de données automatisées pour **collecter, transformer et stocker (ETL)** des données provenant de sources multiples (CRM, bases de données internes).
- Création de **tableaux de bord Power BI** : Développement de rapports et de visualisations interactives pour analyser les **KPI** (chiffre d'affaires, niveaux de stock, satisfaction client) et fournir des insights exploitables aux équipes métier.

Développeur Back-end | Stage | CERIST (5 Mois)

Février 2022 - Juin 2022

- Conception et gestion de la base de données pour stocker les documents et les métadonnées associées.
- Développement du back end (**Django**) avec des **API** pour l'interaction avec la base de données et la gestion des utilisateurs

PROJETS :

Développement et Déploiement d'un Système RAG pour l'Assistance Juridique en droit civil :

- Développement d'un système **RAG** pour l'assistance juridique, intégrant une **API** et une interface utilisateur.
- Utilisation de CamemBERT pour l'encodage des questions, Pinecone pour la recherche de similarité, et Flan-T5 pour la génération de réponses.
- Déploiement complet sur **AWS** avec **Docker**, Amazon ECS, ECR et un pipeline **CI/CD** via GitHub Actions.

Développement et Déploiement d'un Modèle de Reconnaissance de Langage des Signes avec Monitoring :

- Conception et déploiement d'un **pipeline de Machine Learning** conteneurisé avec **Docker**, incluant une **API** de prédiction (**FastAPI**), une interface utilisateur (**Streamlit**), et un système de reporting (**Evidently**) pour surveiller les performances du modèle et détecter les dérives de données.

Mise en œuvre d'un pipeline de données avec Airflow, Spark et DuckDB :

- Orchestration des workflows avec **Apache Airflow**, traitement des données en lot avec **Apache Spark**, et stockage et requêtes efficaces via **DuckDB** pour assurer une gestion optimale des données à grande échelle.

Pipeline d'Analyse, Traitement et Classification de données textuelles (NLP et Machine Learning) :

- Implémenté des techniques avancées de traitement de texte (**Tokenisation, Lemmatisation, TF-IDF, Word2Vec**) pour la normalisation et la vectorisation de données textuelles.
- Comparé des modèles supervisés (**Random Forest, k-NN..**) sur des jeux de données textuelles pour des tâches de classification.

COMPÉTENCES:

- Langages de programmation : Python, R, Java, JS, C.
- Data Engineering: Spark, PL/SQL, ApacheStorm, DAX.
- Outils: Git /Github, Excel, Docker, AWS, Tableau, PowerBI.
- Bibliothèques: Numpy, Pandas, Sckit-Learn, Matplotlib, Pytorch, Tenserflow, OpenCv, Django, Node.js.
- Langues : Français, Anglais.