

Kenzi Lakhdari

Paris, Ile-de-France, France ✉ knzlakhdari@gmail.com ☎ +33769034575



PROFIL

Passionné par les données et leur potentiel, je suis un jeune Data Engineer à la recherche d'une opportunité captivante. D'un naturel sérieux et impliqué, je souhaite consacrer mes compétences à soutenir votre organisation dans la gestion et l'exploitation efficace de ses données.

EXPERIENCE

Data Engineer

Unicancer

January 2023 - September 2023, Paris

• ODH :

- Définir, concevoir et développer un modèle NER (Named Entity Recognition) spécialisé dans l'analyse de comptes rendus médicaux liés aux détections et traitements du cancer. Ce modèle permet d'extraire des informations clés liées au cancer.
- Conception et comparaison de plusieurs modèles (**Spacy, PyTorch, TensorFlow**) via des métriques de confiance.
- Création d'ensembles de test et d'entraînement représentatifs des données médicales pertinentes pour entraîner le modèle NER spécialisé.
- Mise en place d'un serveur hébergé web avec Label Studio pour permettre l'annotation du flux de données par plusieurs équipes (Data Managers & Data Ingénieurs).

• ODH :

- Récupération automatique des données cliniques et médicales pertinentes provenant de différents centres de lutte contre le cancer.
- Restructuration des données à l'aide de **Jupyter Notebook** et **Pandas** pour assurer la cohérence.
- Utilisation d'**Airflow** pour la planification et l'automatisation des workflows, garantissant ainsi une récupération automatique fiable des données.
- Intégration des données dans une base de données centralisée à l'aide de **Cx_Oracle** et **SQL**.

• CANTO :

- Réalisation et maintenance d'un script de migration de données réalisé en **Python**.
- Récupération des données et restructuration de ces dernières pour qu'elles soient adaptées au nouveau modèle logique.
- Utilisation de Docker pour la conteneurisation.

• MyPEBS :

- Traitements de données génomiques à l'aide d'outils Big Data comme Spark.
- Mise en place d'une distribution parallèle à l'aide de **PySpark** pour accélérer le traitement des données.
- Intégration de données massives (Bulk Loading) avec **SQLoader**.

EDUCATION

Ecole d'Ingénieur du Littoral Cote D'Opale Spécialité Informatique

Minor in Data/AI • France • 2020-2023

- Compétences en statistiques et en mathématiques.
- Identifier les objectifs de l'analyse des données avec les managers et chefs de projet.
- Recueillir des données à partir de sources internes et externes (ETL).
- Raisonnement analytique et connaissance des langages de programmation (SQL, Python, **R, Java, Dax**).
- Manipulation de données avec R et Python (Pandas, scikit-learn...)
- Data Viz : Création et analyse de tableaux de bord (**Power BI**, Excel, Matplotlib)
- Déploiement de modèles d'apprentissage automatique (SVM, Xgboost...)
- Utilisation des méthodes agiles pour une gestion de projet flexible et efficace (**Docker, Jenkins**)

CERTIFICATIONS

Artificial Intelligence in Healthcare

Stanford School of Medicine • 2023

SKILLS

Programmation : Python, R, SQL, NoSQL, Java, DAX

Visualisation : Power BI, Dash, Dataiku, QlikView, Matplotlib

Office: Excel, Word, PowerPoint, Teams...

Cloud & Orchestration : AWS, DataBricks, Airflow

Langues : Français Natal, Anglais C1 (TOEIC 975/990);