

## PALE OLLO SALOMON

Data Engineer

Contact :

Adresse : Tunis, Tunisie

Téléphone : +216 21293558

E-mail : ollosalomon@gmail.com

GitHub : <https://github.com/ollosalomon>



---

### Profil Professionnel

Ingénieur en données passionné par l'industrialisation des modèles d'IA et l'optimisation des pipelines de données. Spécialisé dans la conception et le déploiement de solutions scalables et performantes, intégrant les meilleures pratiques de l'ingénierie des données et du MLOps. Expérience avérée en intégration de modèles NLP et en gestion de flux de données massifs avec des outils Big Data et DevOps.

### Expérience Professionnelle

#### Smart Conseil, Tunisie

**Data Engineer** (Janvier 2023 – Présent)

- Développer et déployer des modèles NLP avec TensorFlow, FastAPI et Docker réduisant le temps de traitement des données textuelles.
- Concevoir et maintenir des pipelines de données pour l'extraction et la transformation des données brutes.
- Implémenter une plateforme analytique en temps réel, améliorant l'accessibilité des indicateurs clés pour 10+ utilisateurs quotidiens.
- Déploiement automatisé des projets avec GitLab (CI/CD) et Docker.
- Développer une application d'analyse linguistique avec Python, Django Rest Framework, modèles NLP (TensorFlow, FastAPI), PostgreSQL et Docker analyser 500+ entrées quotidiennes.
- Sécuriser les données et garantir leur intégrité.
- Documentation des développements et audits de qualité de code avec Mkdocs et SonarQube.
- Support technique et optimisation des processus de gestion des données.

### Compétences Techniques

- **Langages** : Python, SQL
- **Big Data & Pipelines** : Hadoop, Kafka, Airflow, ETL, Snowflake, Databricks, PySpark
- **Machine Learning & NLP** : TensorFlow, Scikit-learn, Polars, FastAPI

- **Bases de données** : PostgreSQL, MongoDB
  - **Cloud & DevOps** : Microsoft Fabric, Docker, CI/CD (GitLab), AWS (notions)
  - **Visualisation** : Power BI
  - **Méthodes Agiles** : Scrum
  - **Soft Skills** : Esprit critique, prise de décision efficace, gestion du stress, travail en équipe
- 

## Éducation

- Master de Recherche en Science de l'Informatique - ISIMS, Tunisie (2020 – 2022)
  - Licence Fondamentale en Science de l'Informatique - FSM, Tunisie (2017 – 2020)
- 

## Projets Personnels

### Pipeline ETL d'Analyse des Sentiments sur Twitter ([Lien GitHub](#))

#### Objectifs :

- Collecter des tweets en temps réel.
- Effectuer un prétraitement des données (nettoyage et tokenisation).
- Appliquer un modèle d'analyse des sentiments.
- Agréger les résultats pour créer des rapports et des visualisations interactives avec Power BI.

**Technologies** : Airflow, Docker, Python, Snsrape, NLP, Power BI.

### Pipeline ETL avec PySpark, Airflow et PostgreSQL ([Lien GitHub](#))

#### Objectifs :

- Traitement et nettoyage de données issues de capteurs IoT de véhicules autonomes.
- Automatisation du déploiement via Airflow et Docker, stockage optimisé avec PostgreSQL.

Technologies : PySpark, Apache Airflow, Docker, PostgreSQL, GitHub Actions.

### Analyseur de Sentiments ([Lien GitHub](#))

- Extraction de données via Snsrape, labellisation avec Detoxify, et stockage dans MongoDB.
- Création d'un modèle d'analyse de sentiments et déploiement avec FastAPI et Docker.

### CSV to PostgreSQL ([Lien GitHub](#))

- Importation et sauvegarde de données dans PostgreSQL à partir de fichiers CSV.

**Technologies** : Docker, Mkdocs, Pandas, Python, PostgreSQL.

## Autres Réalisations

Article publié : Arabic Hate Speech Detection System Based on AraBERT

<https://ieeexplore.ieee.org/document/10101577>

---