



Ollo Salomon PALE

DATA ENGINEER JUNIOR / DATA SCIENTIST NLP

☎ 00216 21293558 📞 00216 21293558

✉ ollosalomon@gmail.com

📍 Tunisie, Tunis Nour Jaafar, Cebalat Ben Ammar

<https://github.com/ollosalomon>

<https://www.linkedin.com/in/ollo-salomon-pale-1a7576152>

PROFIL

PROFESSIONNEL

Passionné par l'ingénierie des données et l'IA, je suis un Data Engineer Junior et Data Scientist NLP. Avec une expertise en Python, TensorFlow, et UML, je conçois et évalue des modèles IA avancés. De plus, je développe et déploie des pipelines robustes pour l'acquisition, le nettoyage et la transformation des données provenant de différentes sources. Mon expérience inclut également la gestion d'équipes. Je suis prêt à apporter une approche rigoureuse et créative à tout projet, en alliant passion et expertise technique.

COMPETENCES

-Langage de programmation :

- PYTHON
- C++

-OS :

- WINDOWS
- LINUX

- Outils de développement :

- Visual studio Code
- Jupyter & Jupyter Lab
- SSH

-Base de données :

- Postgres SQL
- SQL

-Méthodologie et Modélisation :

- AGILE
- UML (Power Design)

-Big DATA:

- Kafka
- Snowflake
- SnowSQL
- Data Build Tools (DBT)
- Airflow
- ETL/Data Pipeline

PARCOURS PROFESSIONNEL

SMART CONSEIL, Tunisie, Tunis (<https://smarte-conseil.fr/fr/>)

01/23 - à présent

Rule_management

- Projet où les Product Owners définissent les KPIs.

ML_engine

- Gestion des Règles et modèles d'intelligence artificielle.

Projet 4inshield :

Contexte du projet :

Développement d'une application pour gérer les activités des enfants dans une école.

Tâches effectuées :

- Création d'APIs pour les fonctionnalités de l'application avec Django.
- Création des modèles d'analyse de sentiments pour détecter les sentiments des enfants (TensorFlow, Scikit-learn).
- Définition des KPIs de l'application avec les Product Owner.
- Création de règles basées sur les résultats des modèles, application de ces règles pour détecter les tendances des enfants (agression, intimidation...).
- Mise en place des jobs périodiques avec Celery pour l'application des règles de manière périodique.
- Mise en place d'un système de crawl (Selenium, Python) orchestré par Airflow.
- Sauvegarde des données dans Postgres SQL.
- Envoi des données en temps réel des applications Rule_management, ML_engine, Système de crawl et 4Inshield avec Kafka.
- Documentation du code.
- Versioning du code avec GitLab.
- Utiliser Snowflake comme datawarehouse pour stocker certaines données
- Créer des Dashboards avec Power BI
- Manager les stagiaires (Faire des daily, attribuer certaines tâches de création de modèles aux stagiaires)

REGIM Laboratory, National School of Engineers (ENIS), University of Sfax, Tunisia | Sfax | Stage

02/2022 - 10/2022

Contexte du projet :

Elaboration d'un système de détection et de classification des discours de haine sur les réseaux sociaux à travers le Deep Learning.

-Versioning :

- Github, Gitlab, CI/CD

-Reporting :

- Power BI

-Devops :

- Docker
- Flask (pour le déploiement des modèles IA)

-Framework IA :

- Tensorflow
- Keras
- Scikit-learn

-Framework Dev :

- Django (Très bonne connaissance Back end)
- Streamlit

-Package & package IA :

- Package python (nécessaire)
- Transformers
- Pandas
- Numpy
- Twint, Selenium, snsrape
- MkDocs

- Faire du benchmarking des modèles

- Maîtrise du mécanisme d'attention pour les modèles Deep Learning

- Concevoir et évaluer des modèles Deep Learning/machine Learning

- Gestion des stagiaires Data scientist NLP

PROJETS REALISES

- Deep learning : Mise en place d'un système de détection et de classification des discours de haine sur les réseaux sociaux à travers le deep learning.

- Auteur de l'article: Arabic hate speech detection system based on Arabert

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

AUTRES COMPETENCES

- Esprit d'équipe
- Grande autonomie
- Organisation
- Pianiste/Beatmaker

CENTRES D'INTERET

- Technologie/Recherche

Tâches réalisées :

- Scraper les données avec les APIs Twitter, Twint, snsrape et Selenium.
- Concevoir une base de données et l'annoter.
- Créer des visuels avec Matplotlib et Seaborn.
- Effectuer du feature engineering et du feature scaling (normalisation, standardisation).
- Concevoir et évaluer des modèles de Deep Learning à l'aide de :
- TensorFlow
- Keras
- Optimiser les modèles construits grâce aux techniques de :
- Batch Normalisation
- Réglage des hyperparamètres
- Déployer des modèles de Deep Learning avec le framework Streamlit.

Association Afrique Intelligence | Sfax | Bénévole

Responsable Artistique et culturel

03/2022 - 03/2023

S'informer en permanence sur l'art et la culture afin de l'utiliser pour sensibiliser les populations en Tunisie sur la migration africaine vers l'Europe

ETUDE ET FORMATION

Institut supérieur d'informatique et de multimédia | Sfax

Master de recherche en sciences de l'informatique : Mention Assez Bien
10/2020 - 11/2022

- Master 2 Recherche en sciences de l'informatique

Le Master recherche en Sciences de l'Informatique a pour objectif de fournir une formation approfondie sur la façon de concevoir, d'optimiser et de mettre en œuvre des systèmes complexes et les dernières techniques dans le domaine du data mining.

10/2017 - 07/2020

Faculté des sciences | Monastir

Licence fondamentale en sciences de l'informatique : Mention Assez bien

09/2015 - 07/2016

Lycée Moderne | Agnibilekrou (CÔTE D'IVOIRE)

Baccalauréat : Sciences - Assez bien

LANGUES

Français

Langue maternelle

Anglais

Conversational

