



Abdou Yaya SADIAKHOU

Data Engineer

✉ sadiakhouyaya738@gmail.com

🏠 Chatenay-Malabry (92290)

🌐 <https://abdoufermat-5.netlify.app/>

📞 0760060509

Langues

Anglais

B2

Français

C2

Réseaux sociaux

LinkedIn

[@abdoufermat](#)

Github

[@abdoufermat5](#)

Centres d'intérêt

Football - Basket - Manga/Animes

Certifications

Algorithmic Toolbox

0/1 Knapsack, DP, Tri/Insertion..
([voir](#))

Google Cloud Platform Big Data and Machine Learning Fundamentals

construire un pipeline de données en continu (auto-scaling) pour ingérer, traiter et visualiser les données.
([voir](#))

Mathematics for Machine Learning: Linear Algebra

Implémentation sur python des vecteurs et matrices (vecteurs, valeurs propres..), traitement d'images([voir](#))

Diplômes et Formations

M2 DATASCALE

Depuis septembre 2022

Université Versailles Saint-Quentin Versailles, France

Diplôme d'Ingénieur de Conception

De novembre 2017 à juillet 2022

Ecole Polytechnique deThiès Thiès

Expériences professionnelles

Stage de fin d'études Data Engineer

De mai 2023 à septembre 2023 Banque de France Paris

- **Conception de framework de logging/monitoring/alerting**

- Mise en place d'un standard de logging applicatif sous forme de package python
- Indexation de logs multi-sources sur elasticsearch avec filebeat comme agent
- Mise en place de dashboard de monitoring/alerting avec kibana

- **Devops**

- Mise en place de systèmes d'orchestration de conteneurs docker
- Automatisation de process PostgreSQL
- Tests unitaires et test d'intégration des microservices

- **Microservices**

- Etude et réalisation de microservices **Data Driven (Machine Learning)**
- Décomposition des applications en microservices (refonte architecturale)

- **Réalisation de POC au sein de l'équipe ILIADE**

Outils : Python, Docker (Compose, Swarm), Elasticsearch, Logstash, Kibana, Jira, Jenkins, PostgreSQL

Développeur web mobile

De mai 2021 à septembre 2022 TEAMX GROUP Dakar, Senegal

- Développement android native (Java/Kotlin): mobile banking app ([XaalyS](#))
- Vuejs et Nuxtjs: architecture microfrontend B2B2C app ([Kwely](#))

Projets personnel

Calculs statistiques pour le Cloud Personnel

De novembre 2022 à mars 2023

Université Versailles Saint-Quentin Versailles

- Etude de données GPS (geolife) avec les outils Scikit-Learn et Geopy
- Déterminer les trajets Domicile-Travail
- Application de méthodes heuristique
- Déterminer le type de transport (Marche, Velo, Train, Bus..)

Voir le site: [_gps-streamlit](#)

Projets personnel

● Evaluation de la qualité des données et mapping

De septembre 2022 à décembre 2022

Université Versailles Saint-Quentin Versailles

- Etude des taux de polluants (Paris, Yvelines, Haut de Seine)
- Evaluation facteurs de qualité (complétude, conformité, hétérogénéité, doublons..)
- Calculs de mappings table cible (ville, stations)
- Implémentation sur Talend Data Integration

Compétences

Elasticsearch - Logstash - Kibana

- Utilisation de l'ensemble ELK pour la gestion des logs et l'analyse des données.
- Configuration avancée des pipelines de Logstash pour l'ingestion des logs de diverses sources de données.
- Utilisation d'agent comme filebeat pour exporter les logs vers elasticsearch ou logstash
- Mise en place de template dans kibana et de dashboard pour le monitoring

Python

- Frontend-Backend [Microservices]
 - Développement de services par protocoles: SOAP, gRPC, REST, GraphQL
 - Technos: DjangoREST, Fastapi, Flask, gRPC-Python...
- Traitement distribués et parallèles: Spark (PySpark)
- Développement d'application Streamlit

Cloud et Devops

- Docker
 - Rédaction de dockerfiles pour conteneuriser les applications (orientés data)
 - Orchestration de conteneurs et optimisation de ressources (docker compose)
- Tests unitaires et tests d'intégration
 - Pytest et Unittest pour les tests unitaires
 - Selenium, Postman pour les tests d'intégration
 - Jenkins

Langages de requêtes

- SQL, XQuery, PigLatin... (avec #oracle, #mysql, #postgresql, #hdfs)

Javascript

- développement d'applications web (front/back): NodeJS et VueJS

Hadoop Map Reduce & Spark

- Mise en place de pipeline d'ingestion de données
- Traitement distribués et parallèles

Talend Data Integration

Intégration de données

- Mise en place d'une ETL
- Evaluation de la qualité