



# Frimpong Michel-Ange ADOTRI

## ÉTUDIANT INGÉNIEUR EN BIG DATA ET MACHINE LEARNING

### CONTACT

📍 Noisy-Le-Grand, 93160, FRANCE

✉️ [frimpong.adotri@gmail.com](mailto:frimpong.adotri@gmail.com)

🌐 [Frimpong\\_Michel-Ange ADOTRI](https://www.linkedin.com/in/Frimpong_Michel-Ange_ADOTRI)

☎️ +33 07 53 64 95 57

🐙 <https://github.com/frimpong-adotri-01>

### COMPÉTENCES

- ▶ Langages de programmation : PYTHON, Java, C++, Scala, SQL,...
- ▶ Traitement de données : PANDAS, SCIKIT-LEARN, SEABORN, NUMPY, STREAMLIT, Plotly
- ▶ Big Data: APACHE SPARK, Apache Kafka, Hadoop, HDFS
- ▶ Bases de données: INFLUX-DB, MongoDB, PostgreSQL, Cassandra
- ▶ Logiciels de conteneurisation : DOCKER, Portainer
- ▶ Projet Versioning : GIT, GitHub
- ▶ Cloud: AZURE, aws

### SOFT SKILLS

- ▶ Curieux
- ▶ Proactif
- ▶ À l'aise à l'oral
- ▶ Rigoureux

### LANGUAGES

- 📖 FRANCAIS : LANGUE MATERNELLE
- 📖 ÉWÉ : LANGUE MATERNELLE
- 📖 ANGLAIS : B2

Actuellement élève ingénieur en Big Data et Machine Learning à l'EFREI-Paris, je suis également apprenti Data Engineer chez Informatique-CDC, la filiale IT de la Caisse des Dépôts. J'ai également eu une expérience assez poussée en tant qu'apprenti Data Scientist à Mutex, une entreprise d'assurance. C'est au cours de mon expérience de Data Engineer que j'ai embrassé le métier et développé une passion pour la Data Ingénierie. Ainsi, arrivé en fin de mon Cycle Ingénieur (cycle M), je suis à la recherche d'un contrat à durée indéterminée en tant qu'Ingénieur dans le Big Data et le Machine Learning afin de poursuivre ma progression technique et découvrir de nouvelles facettes du métier.

### EXPÉRIENCE PROFESSIONNELLE

#### Data Engineer

CDC-Informatique | Octobre 2023 - Septembre 2024

Actuellement, je suis Data Engineer en alternance au sein de la Direction des Études Informatique de CDC-Informatique. J'ai principalement:

- Codé un système d'ingestion de fichiers CSV et Excels (deux en un) spécifiques dans un système HDFS [Langage/Technologies: Python, Spark, Hadoop, Hive]
- Mis en place des pipelines transfert de données [Langage/Technologies: Python, Spark, Hadoop, Hive]

#### Data Scientist

MUTEX | Septembre 2022 - Octobre 2023

- Le développement de modèles prédictifs pour la prévoyance collectives
- Le traitement et nettoyage de données, l'automatisation de certains traitements et le développement d'un package Python pour l'équipe Data & Analytics de MUTEX
- La data visualization

### FORMATIONS

#### Cycle ingénieur - parcours Big Data and Machine Learning | 2022 - 2024

EFREI-Paris, Villejuif 94800

Machine Learning, Data Visualization, Data Engineering, Cloud Computing, DevOps, Functional programming with Scala, ...

#### Double licence Mathématiques-Informatique | 2019 - 2023

Université Paris Cité, Paris 75006

### PROJETS

#### Big Data Application Project (sur GitHub)

EFREI-Paris | Janvier 2024

- Mise en place d'une infrastructure mettant en exergue un modèle de machine learning en production (API), un système de monitoring avec Prometheus et Grafana

Langage/Outils : Python, Grafana, Prometheus, Microsoft Azure, Docker

#### Data Engineering Project (sur GitHub)

EFREI-Paris | Mars 2023 - Mai 2023

- Mise en place de l'architecture Docker (Docker Compose)
- Initialisation et configuration d'une stream Kafka
- Envoie des données de la stream Kafka vers S3
- Développement de 2 microservices de Data Visualization et de monitoring et ajout des microservices à l'architecture Docker via Docker Compose

Langage/Outils : Scala, AWS S3, Docker, Python, Streamlit, Apache Kafka, Apache Spark, GitHub Actions

#### Diagnosis of COVID-19 and its clinical spectrum (sur GitHub)

EFREI-Paris | Octobre 2022 - Novembre 2022

- Exploration et analyse des données des patients d'un hôpital de Sao Paulo dans le contexte du COVID-19

Langage/Outils : Python, Streamlit, Plotly, SciKit-Learn

#### Construction d'un système IoT open-source pour serre aquaponique

Université Paris Cité & DeltaLab Prototype | Janvier 2022 - Mai 2022

- Mise en place de l'infrastructure logicielle du système IoT via NodeRED (configuration du broker MQTT et du serveur Docker, etc.)
- Configuration et administration de la base de données chronologique InfluxDB
- Mise en place du Dashboard de visualisation des données collectées
- Documentation du projet (manuels d'installation, spécifications, ...)

Langages/Outils: Docker, InfluxDB, NodeRED, Langage C, CSS, Mosquitto