

# YAHIA AKKAZI

@ yahia.akkazi@centrale-casablanca.ma  
yahia-akkazi yahiaakkazi

+212656614875

Ville verte, Bouskoura

Casablanca, Maroc



## EXPÉRIENCE PROFESSIONNELLE

### Orange

#### Data Engineer/Scientist intern

Mars 2023 - Septembre 2023

Paris

- Traitement de 2.3 milliards d'enregistrements par jour, ingestion, aggrégation & enrichissement de données, entraînement du modèle & déploiement.
  - Mise en place de toute l'architecture des pipelines de données et déploiement de l'application de la prédiction, sur le cloud (GCP).
- Résultat:** Qualité de service améliorée & coûts économisés.

### BNP Paribas

#### Stage Data engineer

Mars 2022 - Septembre 2022

Paris

- Ingestion de données en batch depuis PostgreSQL, vers neo4j en ETL.
  - Design & Implémentation de pipeline de données assurant le traitement & la migration de 800k d'enregistrements par jour.
  - Développement d'un Dashboard de monitoring de la performance de la solution. python logger -> Kafka -> ELK.
  - Mise en place de la chaîne CI/CD pour déployer la solution.
- Résultat:** Migration fiable et traçable de données vers une base graphique.

### CHR ASTER-iT

#### Stage Data / DevOps Engineer

Juin 2021 - Seprembre 2021

Aix-en-Provence

- Mise en place de pipelines d'intégration et de déploiement continues (CI/CD) pour un ensemble de sites web et applications.
- Mise en place de tests unitaires & fonctionnelles, automatisation de déploiement avec Iac (Ansible), et développement en utilisant du BDD et du TDD.

**Résultat:** Amélioration de l'approche DevOps de l'entreprise.

## PROJETS ACADEMIQUES

### Centrale Marseille - TechnipFMC

#### Projet Machine Learning

Octobre 2021 - Mars 2022

Marseille

Prédiction de la durée de vie d'un appareil installé au fond de la mer en se basant sur ses mouvements.

- Analyse statistique des données séries temporelles + Implémentation d'une architecture CNN LSTM pour prédire la durée de vie.

### Centrale Casablanca - RATP Dev

#### Projet Data Science

Novemb 2020 - Janvier 2021

Casablanca

Destiné aux agences de tramway Casablanca, l'objectif du projet proposé par RATP Dev est d'anticiper le besoin d'agents de stations de tramway en prévoyant le nombre de voyageurs par station.

- Visualisation et analyse exploratoire.
- Développement de différents modèles statistiques et de ML (SARIMA, Prophet, forêts...) et réalisation de benchmarks à l'aide des métriques d'évaluation d'erreur adéquates (RMSE, MAE, MAPE...).

## FORMATION

### Ecole Centrale Marseille

Coursus d'ingénieur de données

#### Parcours Data Science

Janvier 2021 - Septem 2023 Marseille - Nice

Optimisation - Aide à la décision - Cloud computing et informatique distribuée - Génie logiciel orienté objet - Traitement d'images.

### Ecole Centrale Casablanca

Coursus d'ingénieur généraliste

#### Option Sciences de données et digitalisation

Septembre 2019 - 2023 Casablanca

Statistiques - Probabilités - Chaînes de Markov - Finance de Marché - Théorie des graphes - Transformation digitale et technologies du Big Data - Machine Learning - Deep Learning - NLP.

### Lycée Al Khansaa

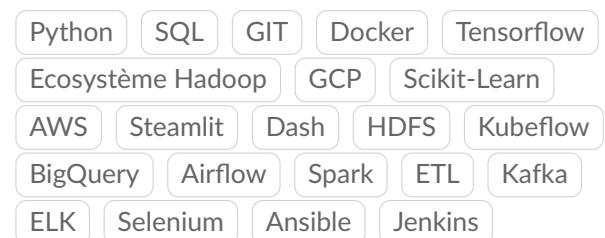
Classes préparatoires MPSI - MP

Septembre 2017 - Juin 2019 Casablanca

## COMPÉTENCES MAITRISÉES

- Data Engineering:** Mise en place de pipelines de traitement de données & gestion du cycle de vie des données.
- Machine Learning:** Analyse exploratoire de données, mise en place d'algorithme de prédiction & déploiement des solutions.

## TECHNOLOGIES MAITRISÉES



## LANGUES

- Arabe : Langue maternelle
- Français : Bilingue
- Anglais : Courant (TOEIC: 930)

## CENTRES D'INTÉRÊT

- Badminton
- Jeux de stratégie
- Sociologie