

Abdou Yaya SADIAKHOU *Data Engineer*

✉ sadiakhouyaya738@gmail.com

☎ 0760060509

📍 Chatenay-Malabry, Hauts-de-Seine

📅 21/09/1998

🚩 Sénégalaise

🔗 my-portfolio

in linkedin

🐙 github



🎓 Education

M2 - DATASCALE: Gestion des données et extraction de connaissances à large échelle,
Université Paris Saclay - UVSQ ☑

09/2022 – present
Versailles, France

Le parcours DataScale se focalise sur l'étude et la maîtrise des nouvelles architectures de gestion de données et d'extraction de connaissances, dont l'échelle varie des très grands serveurs de données (data centers) aux myriades d'objets connectés (edge computing).

Diplôme Ingénieur de Conception, *École Polytechnique de Thiès* ☑

11/2017 – 07/2022
Thiès, Sénégal

L'option Génie Informatique et Télécommunications de l'école polytechnique de Thiès regroupe en son sein les fondamentaux de l'informatique. Il offre la possibilité de se former dans plusieurs domaines de l'informatique (data science, développement web, sécurité, bases de données, modélisation stochastique...) et au bout des 5 ans vous êtes déjà expérimenté en entreprise avec les immersion en cycle spécialisation (3ans).

👛 Professional Experience

VueJs Developer, *Team-X GROUP* ☑

02/2022 – 09/2022
Dakar, Sénégal

J'ai travaillé sur une application de e-commerce adoptant l'architecture micro frontend avec Vue.Js comme technologie frontend et WordPress pour le backend. Je suis principalement sur la partie frontend et l'intégration continue en développant des bouts d'application (totalement découplé).

Android Developer, *Team-X GROUP* ☑

05/2021 – 11/2021
Dakar, Sénégal

J'ai travaillé sur une application de E-Banking Xaalys. J'ai eu à implémenter la fonctionnalité d'éducation financière sur la partie android. J'ai également implémenté l'authentification à deux facteurs sur l'appli.

📄 Certificates

- Algorithmic Toolbox ☑
- Applied Plotting, Charting & Data Representation in Python ☑
- Google Cloud Platform Big Data and Machine Learning Fundamentals ☑
- Mathematics for Machine Learning: Linear Algebra ☑

📁 Projects

Evaluation de la qualité des données et mapping, *Talend data integration*

09/2022 – 12/2022

Dans ce projet nous avons mis en place un ensemble de jobs Talend pour d'une part évaluer la qualité des données à travers quelques facteurs de qualité (conformité, complétude, hétérogénéité, doublons) et d'une autre de calculer un schéma cible (mapping) à partir des expressions de calculs (jointure, union, projection...).

Calculs statistiques 'privacy-by-design' pour le Cloud Personnel,

11/2022 – present

Analyse statistique et machine learning

Etude d'un ensemble de données de 17621 trajectoires avec une distance totale d'environ 1,2 million de kilomètres et une durée totale de plus de 48 000 heures collecté sur 5 ans. L'objectif est de déterminer les parcours domicile-travail et les moyens de transport les plus couramment utilisés pour effectuer ce parcours à partir des données GPS.

Étude des grenouilles de l'Orégon: prédiction de la détectabilité, *Classification supervisée*

09/2022 – 11/2022

L'objectif final de ce travail est d'arriver à prédire selon un certain nombre de critères, si à partir des données collectés par les capteurs placés sur ces grenouilles on parvient lors d'une surveillance en fin de saison à les capturer ou du moins avoir un visuel sur eux.

ANPR: paiement automatique au niveau des péages, *Matlab & Python*

11/2020 – 01/2021

Dans ce projet nous avons réalisé un système de détection des plaques de voitures. Votre plaque est relié à votre compte bancaire et le débit se fait automatiquement au niveau de l'autoroute à péage.

Nous avons entraîné plusieurs modèles de reconnaissance de texte sur les images du dataset (plaques voitures) afin de choisir le meilleur (Random Forest ici).

FerGame: un jeu d'action, *SQL*

12/2019 – 01/2020

Dans ce projet j'ai conçu une base de données pour un jeu de type FPS. L'objectif étant d'implémenter la base et de faire des requêtes de sélection, jointure... et des opérations complexe (vainqueur d'un tournoi, classement par tournoi, niveau d'armes...).

Languages

English (Writing and Speaking) • French (Native)

Skills

Python

Développement web:
Django (+REST), Flask
Big Data & Machine
Learning: Scikit-Learn,
PySpark, Tensorflow
+ Coding Competitions

Hadoop Map Reduce & Spark

Appliquer le paradigme
MR & Spark pour
résoudre des problèmes à
large échelle

Javascript

Développement web:
Vue.js, React.js, Vanilla,
Node.js

Java & Kotlin

Développement mobile:
Android applications
Développement web:
Microservices

Query Languages: SQL, PIG latin, HIVE, PL-SQL

Interaction avec des
bases de données
structurées ou non
structurées

Matlab

Traitement d'image,
machine learning et
interface graphique

Talend

Intégration de données

Agile +

Git, BitBucket, JIRA,
Notion, Asana, Docker