

DIALLO Yousseuf Salian

Concepteur, développeur en science des données
Microsoft Certified



CONTACT



+337 66 66 98 19



diallo.yousseuf.salian@gmail.com



92700, Colombes, île-de-France

FORMATIONS

- ✓ 2020: **FONDAMENTAUX & FULLSTACK EN DATA SCIENCE-**
Jedha Bootcamp
- ✓ 2018: **MASTER MIXTE BIOCHIMIE & BIOLOGIE**
MOLECULAIRE- Lyon 1
- ✓ 2015: **LICENCE SCIENCE DE LA VIE** - Lyon 1

CERTIFICATIONS PRO

- ✓ 2020: Microsoft Certified AI-900 Azure AI Fundamentals
- ✓ 2020: Microsoft Certified AZ-900 Azure Fundamentals
- ✓ 2020: RNCP VI Européen – Concepteur, développeur en Science des Données

EXPÉRIENCES PROFESSIONNELLES

- ✓ 11/2018 - 08/2019: **Consultant Junior / Safir Consulting**

Tâches:

- Développement d'un outil ETL pour la collecte de données
- Analyse exploratoire et descriptive du comportement des clients
- Développement d'algorithmes de machine learning supervisé
- Évaluation de la performance des modèles générés
- Recommandations à l'équipe marketing

Environnement technique:

- Microsoft SSIS, SQL Server 2014

Exploits:

- Développement d'un outil ETL, modèles de classifications robustes

PROFILE

Adepte et passionné de data science, je suis en recherche d'un poste de Data Scientist ou Data analyste au sein du département Data d'une start-up, PME, TPE ou filiale d'un grand groupe.

Suivez-moi sur



COMPÉTENCES

- ✓ Langages de Programmation: **PYTHON, TSQL, DAX.**
- ✓ Machine Learning:
 - Supervisé: **MODÈLES DE RÉGRESSION LINÉAIRES ET LOGISTIQUES**
 - Non Supervisé: **CLUSTERING AVEC PCA OU LDA**
- ✓ Deep Learning: **TENSORFLOW 2.0**
 - Computer Vision: **CNN, GAN**
 - Traitement du langage naturel (NLP): **RNN, LSTM**
- ✓ Statistiques : **DESCRIPTIVES & INFÉRENTIELLES**
- ✓ Traitement des données : **PANDAS, NUMPY, SCIKIT-LEARN**
- ✓ Web Scraping: **CHROMEDRIVER, BEAUTIFULSOUP**
- ✓ Visualisation de données: **POWER BI, TABLEAU DESKTOP, SEABORN, MATPLOTLIB .**
- ✓ Technologies: **SSIS, SQL Server, SSMS, Power BI, Data Modeling.**
- ✓ Technologies Azures: **AZURE SQL DATABASE, AZURE DATA FACTORY, AZURE BLOB STORAGE, AZURE DATABRICKS ETC.**
- ✓ Languages de Communication : **FRANÇAIS(Courant), ANGLAIS (Intermédiaire)**

LOISIRS

- ✓ Lecture: **DÉVELOPPEMENT PERSONNEL**
- ✓ Trading: **CFD, CRYPTOMONNAIES**
- ✓ Sport: **Football**

EXPERIENCES PROFESSIONNELLES

2020

✓ 01/2021: **Projet ADSPY**

Contexte:

- Adspy est la plus grande base de données publicitaire Facebook et Instagram au monde.

Tâches:

Développement d'un connecteur permettant de :

- Se connecter automatiquement à ADSPY
- Saisir les critères de recherche pour les produits gagnants
- Récupérer les informations sur ces produits

Des informations comme le nom, les liens Adspy, Facebook et de la boutique, le nombre de commentaires, de like, de partage sur le produit, ainsi que son prix et la liste des pays où la pub est active.

Environnement technique:

- Python, Webscraping, Chromedriver, Jupyter Notebook, class

✓ 2020: **Projet Instagram Bot**

Contexte:

- Projet personnel de data engineering

Tâches:

Développement d'un robot Instagram permettant de :

- Se connecter à Instagram et de s'authentifier
- Parcourir le profil des utilisateurs et de les Follow
- Liker et collecter leurs posts (optionnel)
- Chargement des données dans une BD NoSQL (mogodb)

Environnement technique:

- Python, Webscraping, Chromedriver, Jupyter Notebook, class

✓ 2020: **Projet ChatBot**

Contexte:

- Dans le cadre de mon projet final chez Jedha Bootcamp

Tâches:

- Développement d'un chatbot pour répondre aux questions fréquemment posées sur le COVID-19
- Source de données OMS

Environnement technique:

- Python, Deep Learning, Tensorflow1.13.2, NLTK, Numpy, Jupyter Notebook, Visual Studio Code

✓ 2020: **Projet Yelp**

Contexte:

- Dans le cadre de projets menés lors de ma formation chez Jedha Bootcamp

- Vais-je aimer ce restaurant ?

Tâches:

- Collecte des données depuis un conteneur de stockage Azure (Azure blob Storage)
- Visualisation des données sur Spark
- Création d'un moteur de recommandation

Environnement technique:

- PySpark, Big Data, Databricks, Azure platform

EXPERIENCES PROFESSIONNELLES

2020

✓ 2020: Projet SMS Spam detector

Contexte:

- Jedha bootcamp Project
- Ce SMS est-il un spam ?

Tâches:

- Développement d'une application web avec flask basée sur le machine learning supervisé
- Cette application permet de détecter la présence de spam dans les messages.

Environnement technique:

- Python, FLASK, Machine learning supervisé, Scikit-learn, Jupyter Notebook

✓ 2020: Projet Real or not ? NLP with disaster tweets

Contexte:

- Jedha bootcamp Project
- Quels tweets relatent de véritable catastrophe ?

Tâches:

- Collecte et stockage des tweets via l'API de kaggle
- Développement d'algorithme de machine learning et de deep learning afin de prédire quels tweets concernent de véritables catastrophes.
- Challenge kaggle

Environnement technique:

- Python, FLASK, NLP, API, Machine learning supervisé,,Tensorflow 2.0, Deep learning, Google colab, laboratory,

✓ 2020: Projet Uber pickups

Contexte:

- Jedha bootcamp Project
- Quel est l'endroit optimal pour un chauffeur Uber ?

Tâches:

- Création d'algorithmes de machine learning non supervisé afin de faire des recommandations aux chauffeurs Uber sur les endroits les plus fréquentés en fonction des heures de la journée.
- Utilisation de la méthode des Kmeans et DBscan pour la création de clusters.

Environnement technique:

- Python, FLASK, Machine learning non supervisé, Clustering, Kmeans, DBscan, Jupyter Notebook

✓ 2020: Projet Fraudulent activities

Contexte:

- Jedha bootcamp Project
- L'entreprise X fait du E-commerce et vend des vêtements faits-main. Le but est de construire un modèle d'apprentissage automatique, permettant de prédire si l'achat d'un vêtement doit être considéré comme une transaction frauduleuse ou non.

Tâches:

- Développement d'un algorithme de machine learning supervisé permettant de prédire la probabilité d'une transaction frauduleuse.

Environnement technique:

- Python, FLASK, Machine learning supervisé, Seaborn, Matplotlib, Scikit-learn, Jupyter Notebook

EXPERIENCES PROFESSIONNELLES

2018 - 2020

✓ 2020: **Projet Conversion rate**

Contexte:

- Jedha bootcamp Project
- Est-ce qu'un utilisateur va-t-il convertir ?

Tâches:

- Création de modèles de machine learning supervisé afin de prédire le taux de conversion d'un client sur le site web.
- Recommandations à l'équipe Produit et Marketing afin qu'ils puissent améliorer ce taux de conversion.

Environnement technique:

- Python, Machine learning supervisé, Scikit-learn, Seaborn, Matplotlib, Jupyter Notebook

✓ 2020: **Projet Speed dating challenge**

Contexte:

- Jedha bootcamp Project
- Comment obtenir un second rdv ?

Tâches:

- Visualisation de le but de comprendre ce qui influe sur l'obtention d'un second rdv dans le cadre d'une rencontre organisée entre personnes.

Environnement technique:

- Python, Machine learning supervisé, Seaborn, Matplotlib, Scikit-learn, Jupyter Notebook

✓ 03/2018 – 08/2018: **Ingénieur de Recherche**

Contexte:

- Stage de fin d'études, Master 2 Biochimie et Biologie moléculaire
- Mise au point d'une méthode expérimentale permettant le diagnostic de certains cas de tumeurs oculaires et orbitaires. Travail en collaboration avec l'Institut d'Ophtalmologie Tropicale d'Afrique (IOTA) et le CHU Point G de Bamako.

Tâches:

- Extraction d'ARNm des échantillons sanguins et de biopsies provenant de patients de l'IOTA.
- Amplification d'ARNm en ADNc par la technique de la RT-PCR en présence de biomarqueurs spécifiques (amorces) aux cas de tumeurs.
- Identification de fusions transcrits et de mutations de gènes par électrophorèse et ou séquençage de l'ADN.

Environnement technique:

- Laboratoire Bactério-virologie de l'Institut National de Recherche en Santé Publique au Mali.