

**Data Scientist | ML Engineer | AI DevOps**

**EXPERIENCES PROFESSIONNELLES**

Depuis 07/2020 <b>Burkina Faso</b>	<p><b>Assistant enseignant en IA</b>, Université Nazi Boni</p> <p>Cours Méthodes Statistique en Big Data, master Science des Données</p> <p>Cours Intelligence Artificielle, master Science des Données</p> <p><b>Fonctions générales :</b></p> <ul style="list-style-type: none"><li>✓ Préparer les supports de cours, les travaux pratiques et les évaluations</li><li>✓ Dispenser les cours à distance et en ligne</li></ul>
Depuis 09/2019 <b>Lyon, France</b>	<p><b>Data Scientist – AI Developer</b>, RCTs</p> <p>Développement et déploiement d'algorithmes liés aux données de la recherche clinique et épidémiologique.</p> <p><b>Fonctions générales :</b></p> <ul style="list-style-type: none"><li>✓ Veille technologique en lien avec les algorithmes et l'Intelligence Artificielle,</li><li>✓ Rechercher et développer des algorithmes et des méthodes de Machine Learning (clustering, NLP, analyse séquentielle, ...) pour des études cliniques et épidémiologiques,</li><li>✓ Rédaction des documents techniques nécessaires à la validation et à la maintenance des algorithmes et des méthodes de Machine Learning,</li><li>✓ Développer des algorithmes et méthodes d'analyse des parcours de santé à partir de données secondaires et de données nationales de santé (SNDS, hub de données de santé),</li><li>✓ Participer à la rédaction ou la révision des SOPs (Standard Operating Procedures) et à la rédaction de plans d'analyse,</li><li>✓ Déployer et mise en production des solutions développées</li></ul>
03/2019 – 08/2019 <b>Lyon, France</b>	<p><b>Data Scientist, Stagiaire</b>, RCTs</p> <ul style="list-style-type: none"><li>✓ Développer d'une plateforme (logiciel) de visualisation des résultats d'analyses statistiques,</li><li>✓ Automatiser des tâches récurrentes d'analyse statistiques et visualisation des résultats</li><li>✓ Définition et développement d'algorithmique et de méthodes d'analyse des parcours de soins en dimension big data (cas de la BD de l'assurance maladie de la France).</li></ul>
10/2018 – 02/2019 <b>Grenoble, France</b>	<p><b>Data Scientist, Stagiaire</b>, Laboratoire Pacte</p> <ul style="list-style-type: none"><li>✓ Analyser les déterminants Socio-démographique et politique de l'intensité de l'activité sur twitter en France.</li><li>✓ Recherchez et développez des modèles de ML pour la prédiction du taux de tweets dans toutes les communes de la France,</li><li>✓ Développer et intégrer des outils automatisés de PNL pour l'analyse des sentiments (SVM, Naïve Bayes, LSTM, etc.) et le traitement de textes.</li><li>✓ Visualisation du taux de tweets par cartographie régionale.</li></ul>
06/2018 – 07/2018 <b>Grenoble, France</b>	<p><b>Data Ingénieur, Stagiaire</b>, Laboratoire Inter-universitaire de psychologie (LIP)</p> <ul style="list-style-type: none"><li>✓ Appliquer le machine learning à la méta-analyse (random forest, analyse factorielle des données mixtes)</li><li>✓ Recherche et développement des modèles d'apprentissage automatique et profond (SVM, CNN, LSTM, GAN, etc.) pour la prédiction de la variation de la répliquabilité entre l'échantillon et le paramètre.</li></ul>

	<ul style="list-style-type: none"> <li>✓ Web scraping (algorithme de collecte des données météorologique)</li> <li>✓ Automatisation des process par application shiny</li> </ul>
10/2017 – 03/2017 <b>Grenoble, France</b>	<p><b>Data Science, stagiaire</b>, Becton and Dickinson Company (BD)</p> <p>Adaptive study design shiny application : développement d'une application shiny de simulation des paramètres de performance lors d'un suivi clinique et d'étude épidémiologique</p> <ul style="list-style-type: none"> <li>✓ Développement applications shiny.</li> <li>✓ Développement d'algorithme de suivi Clinique</li> <li>✓ Méthodes: Pocok, O'Brien and Flemming Design, Wang and Tsiatis Design</li> </ul>
12/2016 – 11/2017 <b>Burkina Faso</b>	<p><b>Data-Manager du projet FBR</b> - Centre Muraz &amp; Banque Mondiale</p> <ul style="list-style-type: none"> <li>✓ Contrôle qualité : programmation des outils de collecte avec Cs pro</li> <li>✓ Contrôle qualité : programmation des batch de contrôle de la conformité des données attendues avec stata</li> <li>✓ Synchronisation : programmation des batch d'envoi de données circuit tablette – serveur local – serveur cloud</li> </ul>
04/2016 – 08/2016 <b>Burkina Faso</b>	<p><b>Statistique – Informatique, Stagiaire</b>, Centre MURAZ Burkina Faso</p> <p>Définition d'une base de sondage par modélisation spatiale</p> <ul style="list-style-type: none"> <li>✓ Géostatistique : modélisation par krigeage</li> <li>✓ Définition d'un plan de sondage du modèle</li> <li>✓ Programmation R et mise en place d'un package R</li> </ul>

## EDUCATIONS & FORMATIONS

---

Novembre 2020	RGPD
Décembre 2019	Introduction to the studies carried out using the Medicare Databases
Novembre 2019	Sample of beneficiaries
Novembre 2019	SNIIRAM /SNDS (French Health National Data System) extraction data for project access
Octobre 2019	SNIIRAM / SNDS (French Health National Data System) architecture and data
August 2019	ICH E3
June 2019	Macro-SAS standard
April 2019	ICH E9
April 2019	ISO 14155 : 2011- Clinical Investigation of medical devices for human subjects
2018 – 2019	<p>Master 2 Statistique and Data Science</p> <p>Université Grenoble Alpes, Grenoble, France</p>
2017 – 2018	<p>Master 1 Statistique and Data Science (mention Bien)</p> <p>Université Grenoble Alpes Grenoble, France</p>
2013 – 2016	<p>Licence en Statistiques et Informatiques (mention Très Bien)</p> <p>Université Nazi Boni, Burkina Faso</p>

## PROJETS PERSONNELS

---

**Github** : vous pouvez consulter mes projets open source et suivre mes d'interventions dans les communautés de développeurs des solutions open source : <https://github.com/armelsoubeiga>

**Web** : vous pouvez connaître mieux sur moi, mes travaux et mes actualités : <https://armelsoubeiga.github.io/>

**Blog** : vous pouvez consulter mes publications liées à l'intelligence artificielle : <http://armelsoubeiga.pythonanywhere.com/>

## COMPETENCES TECHNIQUES

---

### Compétences :

- ✓ Compréhension approfondie des **algorithmes d'apprentissage automatique** (régression, classification, clustering, réduction de dimension, factorisation matricielle, processus gaussiens, optimisation bayésienne).
- ✓ Compréhension approfondie des **techniques d'apprentissage profond** (CNN, LSTM, GAN, Transformer).
- ✓ Bonne compréhension des **algorithmes d'apprentissage par renforcement** (Generalized Policy Iteration, Dynamic Programming, Monte Carlo, etc)
- ✓ Solide expérience de travail avec les outils de modélisation : **TensorFlow, Keras, MLFlow, Scikit-Learn**.
- ✓ Solide expérience dans le **traitement automatique du langage naturel**, en particulier la classification de texte, la modélisation du langage et les tâches de séquence à séquence.
- ✓ Bonne connaissance des techniques de **vision par ordinateur**, en particulier du traitement, la classification et la segmentation d'images, la détection et le suivi d'objets, et la reconnaissance d'actions.
- ✓ Familier avec le Cloud et les outils de DevOps (Git, Docker).
- ✓ Solides compétences en développement logiciel (**Django, Flask, Rshiny**) et en gestion de base de données (**MySQL, Postgres, Oracle**).
- ✓ Bonnes capacités d'analyse et de résolution de problèmes ; enthousiaste à l'idée d'apprendre de nouveaux outils et technologies.

**Langages de programmation** : Python, R, SQL, SAS

**Langues** : Anglais : niveau professionnel, Français : maternelle

## CERTIFICATS

---

Advanced Deep Learning with Keras	05/2020, datacamp.com
Advanced NLP with spaCy	05/2020, datacamp.com
Cluster Analysis in Python	05/2020, datacamp.com
Feature Engineering with PySpark	05/2020, datacamp.com
Introduction to Deep Learning in Python	05/2020, datacamp.com
Introduction to Natural Language Processing in Python	05/2020, datacamp.com
Neural Networks and Deep Learning	05/2020, coursera.org
Sentiment Analysis in Python	05/2020, datacamp.com
Supervised Learning with scikit-learn	05/2020, datacamp.com
Web Scraping in Python	05/2020, datacamp.com
Tensorflow 2.0: Deep Learning and Artificial Intelligence	04/2020, udemy.com