

Angele Blandine Feussi Nguemkam

 Ottawa, ON, Canada

 819-318-3576  angiguems@gmail.com

 [LinkedIn](#) |  [GitHub](#) | [Portfolio](#)

---

## PROFIL PROFESSIONNEL

Data Scientist junior / étudiante finissante en sciences des données appliquées, avec une solide expérience en apprentissage automatique, analyse de données, NLP et qualité des données. Capable de concevoir des pipelines ML complets, d'évaluer et comparer des modèles, et de communiquer des résultats techniques de manière claire et exploitable. Profil bilingue français–anglais, à la recherche **d'un poste junior ou d'un stage en data science / IA appliquée.**

---

## COMPÉTENCES CLÉS

**Programmation :** Python, SQL

**Machine Learning :** Régression logistique, SVM, MLP, Random Forest, XGBoost, validation croisée, SMOTE

**Deep Learning :** Réseaux de neurones (Keras/TensorFlow), autoencoders

**NLP :** Nettoyage textuel, TF-IDF, embeddings, classification de textes

**Analyse & Statistiques :** EDA, statistiques descriptives, ACP/ACM, clustering (K-means)

**Big Data:** Apache Spark (DataFrames, RDD), Hadoop (HDFS)

**Outils :** Jupyter Notebook, Git/GitHub, Docker, Excel analytique, Power BI (bases)

**Langues :** Français (C2) | Anglais (B2)

---

## PROJETS EN SCIENCE DES DONNÉES

**Projet NexaStore – Qualité, analyse et reporting des données (Projet intégré – 3 modules)**

- Audit de la qualité des données et définition d'indicateurs (complétude, cohérence).
- Nettoyage, structuration et création de variables analytiques.
- Analyse exploratoire approfondie et restitution orientée décision.

**Stack :** Python, Pandas, NumPy, Jupyter, Git/GitHub, méthodologie Agile.

**Analyse de données massives – NYC Taxi (Apache Spark & Hadoop)**

- Pipeline Big Data pour l'analyse de plus de 12 millions de trajets de taxis (NYC)
- Ingestion dans HDFS et traitement distribué avec Spark (DataFrames & RDD)
- Nettoyage avancé, suppression de doublons et transformations distribuées
- Exploration des patterns temporels et comportementaux

**Stack:** PySpark, Hadoop (HDFS), Spark SQL, Docker, Linux.

### **Classification de SMS – Détection Spam / Ham (NLP & IA)**

- Développement d'un pipeline complet de classification automatique de messages SMS.
- Prétraitement des textes : nettoyage, tokenisation, suppression des stopwords et stemming.
- Vectorisation TF-IDF et réduction de dimension via autoencoder (Keras).
- Entraînement et comparaison de modèles (Régression logistique, SVM, MLP).
- Évaluation des performances (accuracy, précision, rappel, F1-score, matrice de confusion) et interprétation des résultats.

**Stack:** Python, Scikit-learn, TensorFlow/Keras, NLP, Jupyter.

---

### **EXPÉRIENCE PROFESSIONNELLE**

#### **Opératrice de production — North House Food, Ottawa | 2024 – présent**

- Respect strict des procédures, traçabilité et normes d'hygiène.
- Travail en équipe en environnement cadencé.

#### **Technicienne de laboratoire & Responsable qualité — Fondation sociale Suisse, Cameroun |**

**2021 – 2023**

- Analyses biochimiques et microbiologiques.
  - Mise en place et suivi d'un système qualité (SOP, audits internes).
  - Supervision et formation de stagiaires.
- 

### **FORMATIONS**

#### **Diplôme en sciences des données appliquées — La Cité, Ottawa | 2025 – 2026**

Cours clés : Python, SQL, Machine Learning, Big Data, Azure ML, Spark, Hadoop.

#### **Maîtrise en sciences de la santé (Bactériologie & Virologie) — UCAC, Cameroun | 2020**

---

### **BÉNÉVOLAT**

Participation au diagnostic COVID-19 (tests PCR et analyses) — Centre de recherche pour la santé des Armées | 2020 – 2021

**Références sur demande.**