


RESUME

Angele Blandine Feussi Nguemkam

 Ottawa, ON, Canada

 819-318-3576  angiguems@gmail.com

 LinkedIn | GitHub | Portfolio

PROFESSIONAL SUMMARY

I am a junior Data Scientist and final year Applied Data Science student with strong experience in machine learning, data analysis, NLP, and data quality. Proven ability to design end-to-end ML pipelines, evaluate and compare models, and communicate technical results in a clear and actionable manner. Bilingual French and English, seeking a **junior position or internship in Data Science / Applied AI**.

TECHNICAL SKILLS

Programming: Python, SQL

Machine Learning: Logistic Regression, SVM, MLP, Random Forest, XGBoost, cross-validation, SMOTE

Deep Learning: Neural Networks (Keras/TensorFlow), Autoencoders

NLP: Text cleaning, TF-IDF, embeddings, text classification

Analytics & Statistics: EDA, descriptive statistics, PCA/MCA, clustering (K-means)

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

Tools: Jupyter Notebook, Git/GitHub, Docker, Advanced Excel, Power BI (basic)

DATA SCIENCE PROJECTS

SMS Classification – Spam vs. Ham (NLP & AI)

- Built a complete machine learning pipeline for automated SMS classification.
- Text preprocessing: cleaning, tokenization, stopword removal, stemming.
- TF-IDF vectorization and dimensionality reduction using an **autoencoder (Keras)**.
- Trained and compared multiple models (Logistic Regression, SVM, MLP).
- Model evaluation using accuracy, precision, recall, F1-score, and confusion matrix.

Tech stack: Python, Scikit-learn, TensorFlow/Keras, NLP, Jupyter.

NexaStore Project – Data Quality, Analysis & Reporting

- Conducted data quality audits and defined key quality metrics (completeness, consistency).

- Data cleaning, structuring, and analytical feature engineering.
- In-depth exploration data analysis with decision-oriented interpretation.

Tech stack: Python, Pandas, NumPy, Jupyter, Git/GitHub, Agile methodology.

Big Data Analytics – NYC Taxi (Apache Spark & Hadoop)

- Distributed processing of 12+ million taxi trip records using Apache Spark.
- Data ingestion in HDFS, advanced cleaning, transformations, and large-scale analysis.
- Exploration of temporal and behavioral mobility patterns.

Tech stack: PySpark, Hadoop (HDFS), Spark SQL, Docker, Linux.

PROFESSIONAL EXPERIENCE

Production Operator — North House Food, Ottawa | 2024 – Present

- Strict adherence to procedures, traceability, and hygiene standards.
- Teamwork in a fast-paced production environment.

Laboratory Technician & Quality Lead — Swiss Social Foundation, Cameroon | 2021 – 2023

- Biochemical and microbiological analyses.
 - Implementation and monitoring of laboratory quality systems (SOPs, internal audits).
Supervision and training of interns.
-

EDUCATION

Applied Data Science Diploma — La Cité, Ottawa | 2025 – 2026

Key coursework: Machine Learning, Applied AI, Python, SQL, Big Data, Spark, Azure ML.

MSc in Health Sciences (Bacteriology & Virology) UCAC | 2020

LANGUAGES

French (C2) | English (B2)

References available upon request