



Sadate Yves NANFACK

DATA SCIENTIST / ML ENGINEER

COORDONNEES

4 Avenue Marc Pelegrin, 31400, Toulouse

0760527149

yvesngouateu@gmail.com

Disponible dès maintenant

LinkedIn

LANGUES

Français :
Langue maternelle

Anglais :
Niveau C1

COMPETENCES TRANSVERSES

- Proactivité et créativité dans la résolution de problèmes complexes
- Esprit d'analyse et de synthèse orienté résultats
- Capacité d'adaptation dans un environnement exigeant
- Communication claire et efficace des résultats techniques
- Travail collaboratif en équipes pluridisciplinaires

ASSOCIATIF

Students Ex Machina (SEM) 2024 au sein du GREP-MP (Groupe de Recherche pour l'Éducation et la Prospective, Midi-Pyrénées)

Les SEM sont des étudiants de niveau Licence et Master des écoles et universités d'Occitanie, organisés en groupes **interdisciplinaires** pour confronter, sur le sujet du colloque annuel, des points de vue et des expériences diverses.

Activités :

- Co-organisation du colloque "[Intelligence Artificielle dans tous ses états](#)" du 23 mars 2024.
- Rédaction d'un article : "[Intelligence : artificielle ou pas ?](#)"



PROFIL PROFESSIONNEL

Diplômé en **sciences de la donnée**, alliant background technique en **data science classique**, **Machine/Deep Learning**, et **vision business**. **Passionné** par le développement de **solutions IA innovantes**, notamment en **IA générative** et traitement de données complexes. Je souhaite mettre mes compétences au service des **défis business**.



FORMATION

2023 – 2024 **Mastère Spécialisé : Intelligence Artificielle & Transformation Business**
ISAE SUPAERO, Toulouse.

2021 – 2023 **Master : Intelligence Artificielle & Big Data**
Keyce Informatique & IA, Douala.



EXPERIENCE PROFESSIONNELLE

Mars 2025 - Actuellement

Ingénieur IA/ML, APSIDE, TOULOUSE

- Optimisation du **Module d'Extraction d'Information (IE)** pour l'**automatisation du traitement des Appels d'Offres**.
- Développement d'une **architecture RAG modulaire**
- Conception d'un **framework dual d'extraction** permettant le choix entre **LLMSherpa** et **Unstructured Ingest** avec parsing multiformat (PDF, DOCX, HTML...), et **OCR** en vue de **embedding** et **indexation**, avec amélioration de **85% des performances**
- Implémentation d'un **retrieval hybride** avec Embeddings dense, Sparse et Late interaction et reranking ColBERT pour une amélioration de **+30%** de précision.
- Technologies utilisées : **Python, Unstructured, Poetry, Docker** (pour **Qdrant/Ollama**), **Git**.

Janvier 2024 - Septembre 2024 (9 mois)

Stagiaire IA pour le MBSE, CAPGEMINI ENGINEERING, TOULOUSE - Stage

- Évaluation** et **intégration** des dernières avancées en **IA générative** pour le MBSE
- Conception d'un **RAG** pour la **génération automatisée** de modèles MBSE
- Technologies utilisées : **Anaconda, LLMs** (API : Claude-3.5-Sonnet, Voyage-2-code), **NLP**

2021 – 2023 (2 ans)

DATA & IT PROFESSIONAL, Délégation Générale à la Sécurité Nationale (DGSN) | AEROPORT & PORT DE DOUALA (PISCES SYSTEM) | (CNPTI), Douala - CDI

- Conception d'un modèle **LSTM** pour la prédiction de pannes (-30% de temps d'arrêt)
- Développement d'une solution **Computer Vision** réduisant la fraude de 25%
- Administration et optimisation de base de données avec 99.9% de disponibilité



COMPETENCES TECHNIQUES

MACHINE LEARNING & DEEP LEARNING

- Frameworks: **TensorFlow, PyTorch, Keras, Scikit-Learn, Streamlit**
- Méthodologies : **Classification, Regression, Traitement du langage naturel (NLP), OCR, Réseaux de neurones, IA générative, RAG (Embeddings Reranking)**
- Document Processing** : Extraction données non structurées, documents parsing (PDF, DOCX ..)

BASE DE DONNEES & VISUALISATION

- SQL** : **PostgreSQL**, SQL Server
- NoSQL & MongoDB**

Visualisation: Power BI, Matplotlib

DEVELOPPEMENT & INFRASTRUCTURE

- Langages : **Python, Java**
- Big Data: **Hadoop, Spark, Dask, ETL**
- Cloud** : GCP, AWS, Azure
- DevOps: Docker, **Kubernetes**
GitHub, Git



PROJETS ET COMPETITIONS

KAGGLE CHALLENGE-AIBT Hackathon 2024 Image Retrieval (2024) [kaggle](#)

1ère place sur 8 équipes (**score 0.702** vs baselines **0.298/0.059**) dans un **challenge KAGGLE** de traitement de **séries temporelles** d'images satellites.

HUGGINGFACE – Credit Score Application Machine/Deep Learning (2023) [Hugging Face](#)

Conception et déploiement, web app de **prédiction** de **score de crédit** multi-modèles (conçus par nos soins), avec une **précision de 86%** sur un dataset de **614 profils**.

