

Yves Arthur Negoum

STAGIAIRE EN DATA ENGINEERING

Location: Toulouse, France
| Mobile: (+33)759043831

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

Email: yvesarthur.negoumwouatedem@etu.toulouse-inp.fr

PROFIL

Étudiant en 2^e année à l'ENSEEIH en Sciences du Numérique – filière Big Data et Calcul Haute Performance , admis dans le cadre d'un double diplôme avec l'Ecole Nationale Polytechnique de Yaoundé. Je recherche un stage en data science pour mettre mes compétences techniques et mon sens de l'analyse au service de projets innovants et orientés performance.

FORMATION

Sciences du Numérique - Big Data et HPC (2e année) Toulouse INP - ENSEEIHT	Présent Toulouse, France
Génie Informatique Ecole Nationale Supérieure Polytechnique de Yaoundé	Sep 2023-Sep 2025 Yaoundé, Cameroun
Mathématiques et Sciences Physique - Equivalent Classes Préparatoires Ecole Nationale Supérieure Polytechnique de Yaoundé	Sep 2021-Sep 2023 Yaoundé, Cameroun
BACCALAUREAT Sciences Mathématiques et Physique (Mention Très Bien)	Sep 2020-Sep 2021 Yaoundé, Cameroun

EXPÉRIENCES

Complétude Soutien Scolaire	Depuis Sep 2025 Toulouse, France
• Accompagnement des lycées et collégiens en Mathématiques , Physique et Chimie	
IREX (Institut du Retour d'Expérience) Stage en Intelligence Artificielle	Juin 2025 –Aout 2025 Yaoundé , Cameroun
• Évaluation, test et documentation de différentes solutions d'IA pour identifier les outils les plus performants.	
• Utilisation d'outils de traitement et d'évaluation (Python, notebooks Jupyter, etc.) dans un environnement collaboratif en ligne.	
• Initiation au télé-travail	
• Contribution à la rédaction de rapports techniques et recommandations pour l'adoption de solutions adaptées	

COMPETENCES

Langages Data	:	Python (Pandas, Sklearn, Numpy, tensorflow), SQL, Bash
Bases de données	:	PostgreSQL, MySQL, SQL Server, Optimisation de requêtes
Outils	:	Git, CI/CD, Linux
Soft skills	:	Rigueur, Esprit d'équipe, Vulgarisation technique

PROJETS

Pipeline ML – Prédiction de maladies cardiaques	Python,Numpy,Sklearn,XGBoost
• Conception d'un pipeline complet de traitement des données médicales (collecte, nettoyage, transformation, normalisation).	
• Gestion et intégration de données hétérogènes (âge, pression artérielle, cholestérol, etc.) en assurant la qualité et la cohérence des jeux de données.	
Lupus Classification Project (Deep learning Project)	Python,Numpy,Sklearn, tensorflow
• Gestion complète du cycle de données : collecte, nettoyage, intégration et préparation de jeux de données médicaux	
• Développement d'un modèle de machine learning supervisé à base de réseau de neurones multicouche (MLP) pour la détection du lupus.	

LANGUES

Français	: Natif
Anglais	: B2

CENTRE D'INTERET

- Football
- Lecture
- Run & Fit