

ROSALIE CORINE NKOUNGHAWÉ TOMEYUM

Data Scientist – Machine Learning Engineer – Ingénieure IA

Le Havre (76) – Mobile France entière – (+33) 6 05 61 78 16 – rosaliecorinetomeyum@gmail.com –
<https://github.com/rosalie0309>

PROFIL

Master 2 Informatique – parcours Exploration Informatique des Données et Décisionnel (EIDD), Université Sorbonne Paris Nord, soutenu en septembre 2026. Deux expériences de recherche appliquée en Computer Vision (GREAH) et en NLP (CIRAD), avec une publication scientifique à l'issue. Je conçois des modèles de deep learning de bout en bout : préparation des données, entraînement, évaluation rigoureuse et mise à disposition via une application ou une API. Disponible immédiatement pour un premier poste de Data Scientist, ML Engineer ou Ingénieure IA.

COMPÉTENCES TECHNIQUES

Programmation & données : Python, SQL, pandas, NumPy, Jupyter, PostgreSQL, MySQL, Talend

Machine learning & deep learning : PyTorch, TensorFlow/Keras, scikit-learn, XGBoost, CNN, Transformers

Computer vision & NLP : SegFormer, U-Net, VGG16-FCN8, PSPNet, OpenCV, Hugging Face, SBERT, RoBERTa

Évaluation & fiabilité : validation croisée stratifiée, calibration de seuil, Dice, IoU/mIoU, AUC-ROC/PR, analyse d'erreurs

Développement & outils : Flask, API REST, Git/GitHub, Linux, Power BI, LaTeX

EXPÉRIENCES PROFESSIONNELLES

Ingénieure de recherche IA / Computer Vision (stage) – GREAH, Université Le Havre Normandie **2026**

- Projet européen OCTOPUS (partenaires : University of Hull, Inha University, UIB) – segmentation sémantique d'images sous-marines pour la cartographie de la Posidonia oceanica.
- Comparaison de 4 familles d'architectures (SegFormer B0 – B5, U-Net, VGG16-FCN8, PSPNet) sur PyTorch : meilleur modèle à 96,8 % de justesse et AUC-ROC de 0,996, avec 36× moins de paramètres que la référence de la littérature.
- Protocole d'évaluation fiabilisé : validation croisée stratifiée 5-fold, correction d'une fuite de données dans le choix du seuil, ensembles par vote optimal.
- Développement d'une application Flask d'inférence et d'évaluation (visualisation des prédictions et des métriques côté serveur).

Ingénieure NLP / Data Science (stage) – CIRAD – projet SURSY, Montpellier **2025**

- Pipelines de classification de corpus multilingues pour la surveillance syndromique : TF-IDF, SBERT, Transformers et modèles supervisés (AUC 0,99).
- Amélioration de la qualité des données : clustering UMAP + K-Means et suppression automatisée de 4 214 doublons.
- Co-signature d'un article scientifique accepté pour la CNIA/PFIA 2026.

Développeuse Deep Learning – reconnaissance faciale – World Voice Group **2023**

- Modèle de détection et d'identification de visages, exposé via une API REST Flask.

FORMATION

Master 2 Informatique – parcours EIDD – Université Sorbonne Paris Nord – soutenu sept. 2026 **2024 – 2026**

Sciences de l'ingénieur – Data Science & IA – ENSPD, Douala **2022 – 2024**

Licence Informatique – Sciences des données – Université de Yaoundé I **2019 – 2022**

PUBLICATION & LANGUES

Publication : Article CNIA/PFIA 2026 – classification de corpus pour la surveillance syndromique (projet CIRAD/SURSY).

Mémoire de M2 : Segmentation sémantique par deep learning d'images sous-marines – SegFormer, U-Net, VGG16-FCN8, PSPNet.

Langues : Français (langue maternelle) – Anglais (B2 – IELTS 5.5, Speaking 6.0, Writing 6.0)