MOUHAMADOU LAMINE SOW

lamine.2004pro.sow@gmail.com | Paris, France
LinkedIn | github.com/Lamine2004Sow

AI Engineer – Stage de fin d'études de 6 mois dès mars 2027

Élève-ingénieur en dernière année : **IA générative** (LLM, RAG, agents), **machine learning**, **optimisation** et **mise en production**, avec des modèles évalués avant d'être utilisés.

FORMATION

Diplôme d'ingénieur en informatique (Bac+5) Sept. 2024 – 2027
Sup Galilée, Université Sorbonne Paris Nord *Villetaneuse, France*
Cours principaux : Machine Learning, Deep Learning, NLP, Fouille de texte, Recherche opérationnelle

Cycle préparatoire intégré (MPSI) Sept. 2022 – Mai 2024
ESIGELEC – ESMT *Dakar, Sénégal*

EXPÉRIENCES

Stagiaire Machine Learning / Data Science Mai 2026 – Juil. 2026
Sonatel (groupe Orange) *Dakar, Sénégal*

- Développement d'un modèle XGBoost prédisant les risques sur le réseau fibre à 6 mois : **88 % de faux négatifs en moins** que la règle métier en place (PR-AUC 0,81 contre 0,28).
- Prédictions calibrées et **expliquées (SHAP)**, agrégées en cartes de risque par zone (3 à 12 mois) pour aider les équipes opérationnelles à prioriser leurs interventions.

Stage d'observation – Numérique et DevOps Juil. 2024 – Sept. 2024
Sonatel (groupe Orange) *Dakar, Sénégal*
Rédaction d'un état de l'art sur l'automatisation du DevOps par l'IA ; initiation à Angular, LoopBack et MongoDB.

PROJETS TECHNIQUES

Agent LLM juridique : garde-fous, citations vérifiées et évaluation automatique 2026
Python, LangGraph, LLM (DeepSeek), FAISS, BM25, embeddings e5, RAGAS, Streamlit, pytest, Docker

- **RAG de bout en bout** sur environ 5 100 articles de loi : ingestion (nettoyage, dédoublonnage, découpage), recherche hybride FAISS + BM25 (fusion RRF).
- Agent LangGraph : reformulation, **garde-fou** qui traite la question comme une donnée (100 % des 10 questions hors sujet détectées, dont une tentative de **prompt injection**), **vérification déterministe des citations** (92,3 % de précision).
- **Golden set** de 60 questions étiquetées : métriques déterministes, juge LLM, RAGAS ; **test de non-régression** qui échoue si un prompt change sans rejouer l'évaluation. Recherche 80 % → 90 %, 100 % de refus corrects.

Plateforme MLOps auto-réparatrice sur Kubernetes local 2026
Python, FastAPI, MLflow, Docker, Kubernetes (kind), Prometheus, Grafana, Evidently, GitHub Actions

- **API backend FastAPI** servant le modèle sur Kubernetes local ; **CI/CD** GitHub Actions qui échoue si le modèle ne bat pas sa baseline ; Docker Compose.
- **Observabilité** Prometheus et Grafana : alerte sur dégradation → réentraînement automatique, nouveau modèle promu seulement s'il est meilleur de 2 % sur des données inédites (MAE 4,49 → 3,09 lors du choc COVID).

Fine-tuning d'un modèle d'embeddings juridique (en cours) 2026
Python, LLM (DeepSeek), sentence-transformers, multilingual-e5-base, PyTorch, MLflow

- 11 088 paires question/article **générées par LLM** puis filtrées (règles et juge LLM) sur 12 849 ; jeu de test de 125 questions figé avant l'entraînement ; fine-tuning sur CPU, suivi MLflow.

Challenge ROADEF/EURO 2020 : planification de la maintenance du réseau électrique (RTE) 2026
Python, Gurobi, PLNE, recuit simulé

- **Optimisation sous incertitude** : recuit simulé à 0,52 % de la meilleure solution connue ; PLNE Gurobi : 8 optimums prouvés sur 15.

COMPÉTENCES

IA générative : LLM, RAG, LangGraph, embeddings (e5, sentence-transformers), FAISS, BM25, fine-tuning, LLM-as-a-judge, RAGAS
ML et optimisation : PyTorch, scikit-learn, XGBoost, SHAP, MLflow ; PLNE (Gurobi), métaheuristiques
Développement : Python, FastAPI, Git, CI/CD (GitHub Actions), Docker, Kubernetes, Grafana, SQL ; Claude Code, Codex, Antigravity
Langues : Français (natif), Anglais (B2, TOEIC 802) | **Certification :** IBM AI Engineering (Coursera, en cours)