

Amadou MAMANE

AI Engineer | Data Scientist

Vincennes, France | +336.52.78.18.48 | mamanesarki@yahoo.fr | Nationalité française
<https://fr.linkedin.com/in/amadoumamane> | <https://github.com/AmadouMamane>

Spécialiste en intelligence artificielle avec **9 ans d'expérience**, je conçois et industrialise des solutions big data et IA. Mon expertise couvre la **vision par ordinateur**, les **systèmes de RAGs** et l'**IA générative**.

COMPÉTENCES FONCTIONNELLES

Statistiques descriptives et inférentielles (estimation, tests d'hypothèse), **séries temporelles**, **NLP**, **RAG**, **prompt engineering**, vision par ordinateur, deep learning, systèmes de recommandation, IA générative, apprentissage non supervisé, segmentation, **méthodologie agile**, data engineering, **MLOps**.

COMPÉTENCES TECHNIQUES

Frameworks ML: Scikit-learn, Spark ML, TensorFlow, Pytorch, Vertex AI, Statsmodels. **Visualisation de données**: Matplotlib, Seaborn, Tableau, Grafana. **Langages de programmation**: Python, Numpy, Pandas, SQL. **Bases de données**: Hive, BigQuery, Pinecone, Cassandra, Oracle. **Plateformes big data**: Hadoop, Spark, Kafka. **Cloud computing**: AWS, GCP. **Outils MLOps**: Git, GitLab CI/CD, Kubeflow, Airflow, Docker, Kubernetes. **IA Générative**: GPT, Llama, Gemini, Gemma, LangChain, LangGraph, HuggingFace.

PROJETS

Janvier 2024 – Aujourd’hui

Détection du cancer de la peau - Compétition Kaggle ISIC-2024 - [Github](#)

Développement d’une solution de détection du cancer de la peau à l’aide d’**images** et de **métadonnées tabulaires**

- **Feature engineering** sur les données tabulaires.
- Gestion du déséquilibre des classes: **Augmentation** des images, **recherche et intégration de données externes**, application de SMOTE sur les données tabulaires.
- **Fine tuning** de plusieurs **modèles de base**: VGG19, Resnet50, Resnet152V2, Densenet201.
- Création de modèles avancés à l’aide de techniques de **validation croisée** et d’assemblage de modèles simples
- Implémentation/adaptation de **fonctions de coût** et de **métriques de performance**: Focal loss, **weighted binary cross-entropy**, class balanced loss, **Partial AUC**, F1-score, recall.
- Optimisation de la performance d’exécution sur CPU, **GPU** et TPU.

Agent conversationnel interactif - [Github](#) | [Demo](#)

Création d’une application de commande pour restaurant intégrant une interface de discussion en langage naturel, permettant aux utilisateurs de passer commande de manière intuitive. Les principales fonctionnalités incluent :

- **Boucles conversationnelles dynamiques**: interaction continue entre le client, le chatbot et les agents sollicités.
- **Menu dynamique**: présentation du menu du restaurant et des options disponibles en temps réel.
- **Assistance client**: ajustement de quantités/préférences, simulation de transmission de la commande en cuisine.

Système de recommandation de musique - [Github](#) | [Demo](#)

- Construction d’un système de recommandation de musique utilisant les données d’interaction utilisateur, de retour implicite (Implicit feedback) et des chansons.
- Utilisation de techniques de: Rang/popularité, Clustering, Filtrage collaboratif, Factorisation de matrice.
- Implémentation de la solution finale sous forme de système de recommandation hybride utilisant à la fois les méthodes de remplissage de matrice et de filtrage collaboratif à base de contenu.

Environnement technique Scikit-learn, Tensorflow, Surprise, LightFm, Imblearn, Nltk, LangGraph, API Gemini

Environnement fonctionnel Santé, Musique numérique, E-commerce

EXPERIENCES PROFESSIONNELLES

Senior Data Scientist

Paris, France

Cartes Bancaires

Septembre 2023 – Novembre 2023

- Développement d’une solution de **détection d’attaques** sur les commerçants à l’aide de séries temporelles
 - Implémentation d’un modèle d’apprentissage supervisé à base de **gradient boosting** prédisant un montant total d’attaques estimé à 500 kilos euros/semaine pour une précision réglée à 0.6
 - Déploiement sur une infrastructure cloud sécurisée permettant de gérer de grands volumes de données en temps réel tout en assurant la conformité réglementaire.
- **Recommandations** de bonnes pratiques de développement **Apache Spark**

Environnement technique PySpark, Spark ML, Python, Hive, Impala, Oozie, Git, Séries temporelles, Data leakage

Environnement fonctionnel Banque, cybersécurité

Senior Data Scientist

RCI Bank&Services

Paris, France

Avril 2019 – Août 2023

- Conception et déploiement d'un modèle de scoring de l'attrition client, atteignant un score F1 supérieur à 97 %.
- Développement d'une application de **déduplication des données clients avec des techniques de NLP**.
- Participation à la **migration des applications** existantes d'un cluster On Premise vers **Google Cloud Platform**.
- Conduite de plusieurs projets **preuves de concept (POCs)** pour évaluer et intégrer de nouveaux outils.

Environnement technique

Scikit-learn, Spark ML, Scala, Python, GraphFrames, GCP, Nifi, Gitlab CI/CD, Jira, NLP

Environnement fonctionnel

Banque, risque et crédit, connaissance client, méthodologie agile

Data Engineer

Altran | mission Orange

Arcueil Cachan, France

Mai 2017 – Mars 2019

- Développement d'applications pour le **traitement en temps réel** des données 2G, 3G et 4G, avec centralisation de ces données dans un format standard accessible à l'ensemble du groupe.
- Participation à la création de prototypes, aux tests et à la sélection des composants des applications.
- Supervision du réseau mobile avec des **KPIs métiers**. **Mise en production et supervision** des applications.

Environnement technique

Kafka, Flink, Spark, Hadoop, InfluxDB, Grafana, Streamsets, Scala, Java, Git, shell

Environnement fonctionnel

Télécommunications, connaissance client

Data Scientist

Altran | mission SNCF

La Défense, France

Septembre 2016 – Avril 2017

- Mise en place d'applications d'IA permettant d'optimiser le remplissage des trains
- **Collecte des données** de recherches en ligne sur les trains
- Développement d'un **modèle** de prédiction de remplissage de la première classe
- Contribution à l'**amélioration de la qualité de code** et **mise en place de bonnes pratiques de développement**

Environnement technique

Scikit-learn, PySpark, Python, Pandas, Numpy, scala, Spark, Git, Bokeh, Jenkins, Luigi

Environnement fonctionnel

Ferroviaire, yield management, méthodologie agile

Data Scientist

Altran | mission Orange

Guyancourt, France

Novembre 2015 – Septembre 2016

- Participation à l'implémentation de la solution **Customer Experience Management On Demand (CEMOD)** de Nokia permettant l'**amélioration de la performance du réseau mobile** et de la **connaissance client**.
- Mise en forme des données réseaux, facturation et CRM. Intégration des données dans le CEMOD et calcul de plusieurs **scores** pour évaluer la satisfaction client ainsi que la qualité du réseau dans les zones de vie du client.

Environnement technique

Hive, Pig, Java, Oozie, shell, svn

Environnement fonctionnel

Télécommunications, connaissance client

Data Scientist

Société Générale | Stage

Guyancourt, France

Avril 2015 – Septembre 2015

- Implémentation d'un **système de détection d'attaques par déni de service** sur des **logs webs**: Recherche sur l'état de l'art de l'analyse de logs webs. Nettoyage des données et reconstruction des sessions de navigation. **Feature engineering** et **visualisation des données**. Clustering pour identifier les groupes de sessions de navigation. Développement de plusieurs modèles de **classification** pour identifier les robots et les humains.

Environnement technique

Hadoop, Scikit-learn, Pig, Hive, Python, Java, Pandas, Numpy, Scipy, Matplotlib, Git

Environnement fonctionnel

Banque, cybersécurité

FORMATION

Applied Data Science Program: Leveraging AI for Effective Decision-Making

En ligne

Massachusetts Institute of Technology

Janvier - Juillet 2024

MS Ingénierie des Systèmes Informatiques Ouverts

Chatenay Malabry, France

Ecole Centrale Paris

2014 - 2015

Master Recherche Génie Industriel et Informatique

Saint-Etienne, France

Ecole des Mines de Saint-Etienne

2011 - 2012

Ingénieur en Génie Industriel

Tunis, Tunisie

Ecole Nationale d'ingénieurs de Tunis

2009 - 2011

CERTIFICATIONS

- C# developer
- **Applied Data Science Program, MIT**

LANGUES

Français - Langue maternelle. **Anglais** - Professionnel