

JAYA SAI KISHORE NEERUKONDA

Ingénieur Machine Learning | Data Scientist

jaysaishoren@gmail.com | +33 745324032 | [LinkedIn](#) | [GitHub](#) | Nice, France

PROFIL

Ingénieur en machine learning avec 3.5 ans d'expérience professionnelle et un Master en Data Science & Intelligence Artificielle. Expérimenté dans la conception et la mise en production de systèmes de machine learning, avec une solide expertise en Python, TensorFlow, PyTorch, MLOps, CI/CD et le déploiement de modèles, y compris les LLMs et les pipelines RAG. Capacité démontrée à développer des workflows ML évolutifs et fiables, générant un impact métier mesurable.

COMPÉTENCES TECHNIQUES

Programmation: Python, SQL, Pandas, NumPy.

Machine Learning & IA générative: Machine Learning, Deep Learning, Scikit-learn, TensorFlow, PyTorch, XGBoost, Computer Vision, NLP, Fine-tuning, HuggingFace Transformers, LLMs, RAG Pipelines, LangChain, bases de données vectorielles (FAISS, ChromaDB).

MLOps & outils de développement: Git, GitHub, CI/CD, Docker, Jenkins, Ansible, AWS (EC2, S3), REST APIs, déploiement de modèles, automatisation de pipelines, DVC, MLflow.

Bases de données & APIs: MySQL, PostgreSQL, RESTful APIs, FastAPI, Flask.

EXPÉRIENCE PROFESSIONNELLE

Ingénieur Machine Learning – Starlite Infotech Ltd., Bengaluru

Jan 2023 – Sept 2024

- Conception et mise en œuvre de pipelines de machine learning et de deep learning de bout en bout, incluant des workflows NLP, permettant d'améliorer la précision des modèles de 20 à 28% grâce au prétraitement des données, à l'ingénierie des caractéristiques, aux embeddings et à l'optimisation des hyperparamètres.
- Développement et entraînement de modèles de machine learning et de deep learning (réseaux de neurones avec TensorFlow/PyTorch) pour des tâches de classification de texte, similarité et analytique prédictive, réduisant les taux d'erreur de 18 à 22% par rapport aux approches de référence.
- Déploiement de modèles ML/DL et NLP prêts pour la production via des API REST et des environnements cloud, réduisant la latence d'inférence et les efforts manuels de traitement du texte de 20 à 30% grâce à l'automatisation.

Ingénieur logiciel senior – Capgemini Technologies Pvt. Ltd., Hyderabad

Déc 2021 – Déc 2022

- Optimisation des pipelines de machine learning et de données en automatisant le prétraitement, l'ingénierie des features et la validation, améliorant significativement les performances et la fiabilité des workflows.
- Déploiement de services ML conteneurisés à l'aide de Docker, Jenkins et Ansible, réduisant le temps de déploiement de 40% et permettant le serving de modèles via des API REST scalables.
- Mise en place de pipelines de monitoring et de CI/CD de bout en bout, incluant la journalisation, les métriques et le versioning des artefacts, augmentant la fiabilité des modèles de 30%.

Stagiaire en data science – Digital Lync, Hyderabad

Juin 2019 – Avr 2020

- Rationalisation des workflows de préparation ML en structurant les étapes de prétraitement des données, d'extraction de features et de validation, réduisant le temps de préparation de 25%.
- Configuration et optimisation des modèles de machine learning à l'aide de validations croisées structurées et d'optimisation d'hyperparamètres sur plus de 20 expérimentations, améliorant la précision des modèles de 18%.
- Évaluation du comportement des modèles via l'analyse des résidus, l'attribution des features et des diagnostics de variance, réduisant le surapprentissage de 20 % et améliorant la capacité de généralisation.

PROJETS

Système de prévision de la qualité de l'air – Côte d'Azur, France

Python, Scikit-learn, modèles de séries temporelles, LSTM/GRU, MLflow, FastAPI, Streamlit, AWS

- Traitement et structuration de données de séries temporelles de qualité de l'air multi-stations issues du portail **Export Avancé de Geod'air**, incluant le nettoyage des données, l'alignement temporel et la préparation des features pour la prévision.
- Conception et évaluation de modèles de machine learning et de deep learning (Random Forest, XGBoost, LSTM/GRU), avec suivi des expérimentations, des métriques et des artefacts à l'aide de MLflow et DVC afin de garantir la reproductibilité.
- Déploiement du pipeline de prévision sur AWS avec FastAPI pour une inférence scalable, et développement d'un tableau de bord Streamlit pour la visualisation des tendances et des prévisions en temps réel (48h → 24h).

Système de recherche et de résumé d'avis basé sur les émotions

ResNet-50, FAISS, LangChain, Flan-T5, Streamlit

- Développement d'un système d'intelligence artificielle permettant de détecter les émotions faciales à partir d'images et de récupérer des avis clients alignés avec le contexte émotionnel identifié.
- Fine-tuning d'un modèle ResNet-50 sur le jeu de données FER-2013 afin de classifier sept émotions faciales, et conception d'un pipeline efficace de recherche d'avis basé sur FAISS et sentence-transformers.
- Génération de résumés concis à l'aide de Flan-T5 et réalisation d'analyses de sentiment avec un modèle RoBERTa afin d'améliorer l'interprétabilité et l'extraction d'insights.

FORMATION

Data Science Tech Institute, Sophia Antipolis, France

Master en Data Science & Intelligence Artificielle

Oct 2024 – Nov 2026

Gudlavalleru Engineering College, Andhra Pradesh, India

Licence en électronique et télécommunications

Juin 2015 - Avr 2019

INFORMATIONS COMPLÉMENTAIRES

Langues: Anglais (C2 – bilingue) | Français (A2, en cours de B1)

Certifications: Neo4j Certified Professional (2025)

Open-Source Contributions: Contributeur Hacktoberfest 2025 (plusieurs pull requests acceptées)