



Salma OUARDI

ML Engineer

ML Engineer with strong problem-solving skills and a knack for bridging the gap between business and technical needs. Seeking a collaborative team where I can contribute to impactful projects and grow professionally.

✉ salmaouardi1998@gmail.com

☎ +33 758224169

📍 Paris, France

🌐 www.salmaouardi.com

WORK EXPERIENCE (2 YEARS)

AI Engineer (1 year)

Ryte AI - Healthcare & Biotech

09/2023 - Present

Paris, France

Built RyteGPT, an internal LLM-powered assistant for structuring and extracting medical information from patient queries and healthcare data.

Achievements/Tasks

- Built and deployed core NLP modules: Roberta-based intent classifier (Vertex AI), multi-entity extraction with LLaMA 3/Mistral 7B (spaCy + IOB tagging), and BioLORD-based entity linking with cross-encoder reranking.
- Deployed quantized LLMs (AWQ) on Azure GPU VMs with vLLM and multithreaded inference; tracked experiments using MLflow and managed CI/CD with Azure DevOps.
- Generated 1,000+ synthetic medical queries using GPT-3.5/4; collaborated with MDs for annotation (Prodigy) and built F1-based evaluation pipelines for entity/token spans.
- Rebuilt entity resolution pipeline using TF-IDF and fuzzy matching, improving F1 from 77% to 98% on large-scale healthcare datasets.
- Developed Spark pipelines (Scala UDFs) on Azure Databricks for 10+ TB of healthcare data; integrated APIs and normalized data from PubMed, CMS, AHA, and others.
- Maintained robust production quality via test-driven development (PyTest, Pydantic), achieving 100% test coverage and schema enforcement.

Data Scientist GCP | End of Study Internship (6 mos)

Orange

03/2023 - Present

Paris, France

Achievements/Tasks

- Built semi-supervised classification models using Active Learning, K-means, and AutoML to label 40,000+ user comments with minimal manual effort.
- Leveraged BigQuery ML and Google Cloud APIs to scale data preprocessing, clustering, and model deployment across the cloud pipeline.
- Designed and implemented testing protocols for MLOps monitoring, including drift detection and data integrity checks for real-time feedback loops.
- Contributed to aligning machine learning work with product use cases, improving interpretability and model relevance for business teams.

Contact : EOCK BELINGA Benoit - benoit.eockbelinga@orange.com

Data Engineer | End of Study Internship (6 mos)

Stellantis

03/2022 - 08/2022

Casablanca, Morocco

Achievements/Tasks

- Developed PySpark and Apache Airflow ETL pipelines to automate the Carflow MEA Dashboards project and enable reliable daily reporting.
- Optimized Oracle Exadata-based workflows for high-volume data ingestion, reducing latency by 46ms and cutting failure rates by 82%.
- Created monitoring scripts and unit-tested pipelines for data consistency, helping ensure robustness in production.
- Coordinated with cross-functional teams to integrate upstream data sources and align metrics with business KPIs.

TECHNICAL SKILLS

Languages

Python (NumPy, Pandas, Scikit-learn, spaCy, PySpark), SQL, Bash

Machine Learning & NLP

Transformers, Token Classification, Entity Resolution & Linking, IOB Tagging, AutoML, Clustering, PyTorch, LLaMA, Mistral, Roberta, GPT-3.5/4, BioLORD

MLOps & Deployment

MLflow, DVC, FastAPI, vLLM, Model Quantization (AWQ), CI/CD, Docker

Cloud & Big Data

Microsoft Azure, GCP, Apache Spark, Airflow

Tools & Version Control

Git, GitHub, Jupyter, VS Code

EDUCATION

Master of Science, Artificial Intelligence

Paris-Saclay University

09/2022 - 09/2023

Paris, France

Engineering degree | Data and Knowledge

Ecole des Sciences de l'Information

09/2018 - 06/2022

Rabat, Morocco

PERSONAL PROJECTS

CV project: Bias Mitigation For Age Detection (10/2022)

- Estimating the age from images while tackling the bias with respect to the protected attributes (Age, Gender, Ethnicity, Face Expression).
- Data Augmentation using Albumentations and Manual Augmentation using OpenCV. Create a customized loss & Change of base model (NASnet, RESnet)

NLP project: Wav2Vec2 with Hugging Face Transformers (01/2023 - 02/2023)

- Fine-tuned a pre-trained Wav2Vec2 model for Automatic Speech Recognition (ASR), achieving a word error rate (WER) of less than 5% on the test sets.

LANGUAGES

English

Full Professional Proficiency

French

Full Professional Proficiency