



Salma OUARDI

Data Scientist

Data Scientist specializing in AI and ML, 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

☎ 0758224169

📍 Paris, France

🌐 www.salmaouardi.com

WORK EXPERIENCE (2 YEARS)

Data Scientist (1 year)

Ryte AI - Healthcare Startup

09/2023 - Present

Paris, France

Achievements/Tasks

- Development of an open-source LLM-based solution for generating biographies of Health Care Practitioners.
- Enhanced model accuracy with LightGBM and Hyperopt, achieving significant improvements over baseline performance.
- Led the design and implementation of entity resolution pipelines, reducing processing time by 96% and boosting F1-Score from 77% to 98%.
- Deployed scalable data solutions on Azure, leveraging Databricks, VMs, and Azure Storage to enhance processing capabilities. Engineered Spark-based data transformation systems for datasets over 10 TB, utilizing Scala UDFs for complex processing.
- Integrated diverse healthcare data sources, such as Pubmed and Clinical Trials, into the analytics ecosystem, engineering solutions to link doctors and hospitals using data from AHA, CMS, and LexisNexis.
- Led team initiatives by spearheading research and pilot projects to drive adoption of new methodologies and technologies.

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

Orange

03/2023 - Present

Paris, France

Achievements/Tasks

- Aligned data science projects with corporate strategy, employing data preprocessing techniques to enhance machine interpretability and applied ML techniques like Active Learning, K-means, and semi-supervised learning for large-scale data analysis and predictive analytics.
- Labeled 40,000 comments using Active Learning and employed various machine learning tools including AutoML, BigQuery ML, and custom models to facilitate semi-supervised classification.
- Developed and refined testing protocols for MLOps monitoring implementation, integrating Google Cloud APIs for efficient data clustering and labeling.

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

- Led the Carflow MEA Dashboards project, directing key initiatives in big data management and implementing Oracle Exadata for large-scale, high-performance database management, while automating ETL processes to enhance efficiency and reduce errors.
- Developed and maintained robust data pipelines using PySpark and Apache Airflow, ensuring data integrity through rigorous testing and comprehensive documentation of ETL processes.
- Significantly boosted system efficiency, reducing latency by 46ms and failure rate by 82%, through targeted system enhancements and optimization efforts.

TECHNICAL SKILLS

Languages

Python (NumPy, Pandas, Scikit-learn, spaCy), SQL and NoSQL, Spark, Bash

Machine Learning

Supervised & Unsupervised Learning, AutoML

Deep Learning & AI

Transformers, LLMs, PyTorch, TensorFlow, Hugging Face.

Cloud & DevOps

Microsoft Azure, GCP, Docker, CI/CD

Tools & Platforms

Git, GitHub, MLflow, DVC.

EDUCATION

Master of Science, Artificial Intelligence

Paris-Saclay University

09/2022 - Present

Paris, France

Engineering degree | Data and Knowledge

Ecole des Sciences de l'Information

09/2018 - 06/2022

Rabat, Morocco

PERSONAL PROJECTS

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.

Bias Mitigation For Age Detection (10/2022 - 11/2022)

- Developed a model to estimate age from images, addressing bias related to age, gender, ethnicity, and facial expression. Techniques used included data augmentation and custom loss functions.

LANGUAGES

English

Full Professional Proficiency

French

Full Professional Proficiency