

Saad TAZROUTE

Data Engineer & Cloud Engineer @ThePlaceToWish : Ecole Centrale Marseille/Aix-Marseille University looking for new opportunities to shine, available to start from December 2024.

10 Rue Pierre Chalnot, Nancy, France

☎ 0616569063 | ✉ Tazroute.Saad@gmail.com | 🌐 saadtazroute | 🌐 SaadTazroute



Education

- Ecole Centrale Marseille (ECM) : Digital.e, Exchange Program** :Machine Learning and Large Scale Processing, 2021-2022 Advanced Mathematics for Generative AI & Quantitative Marketing, Data analytics (Dataframes and Time series manipulation and analysis). *Marseille,France*
- Aix-Marseille University , Msc. IAAA** : Studied courses : In-depth Machine Learning & Big Data, Deep learning & 2021-2022 Reinforcement Learning, Latent variable models and Bayesian theory, Parametric and Non-parametric estimation, Data mining and visualization, Applied statistics. *Marseille,France*
- Ecole Centrale Casablanca (ECC)**, Engineering program (Bac+5) : Followed courses : Probability, statistics, 2018-2021 optimization, advanced Python, project management, economics, image processing, data visualization ... *Casablanca, Morocco*

Professional experiences

ThePlaceToWish

Nancy, France

DATA & CLOUD ENGINEER

October 2023 – Present

- Led the migration of an e-commerce recommendation system to AWS, implementing a **robust microservices** architecture and **scalable data pipeline**
- Orchestrated the AWS cloud infrastructure using **Terraform**, implementing Infrastructure as Code (**IaC**) for consistent, staged deployments
- Developed an **end-to-end data ingestion pipeline** using AWS Glue, facilitating efficient ETL processes for large-scale data processing
- Implemented **serverless APIs** using AWS API Gateway and Lambda, integrating seamlessly with the e-commerce platform.
- Designed and maintained a high-performance data warehouse solution using Amazon Redshift, enabling fast and complex analytical queries
- Leveraged Amazon Managed Workflows for Apache Airflow (MWAA) to **develop and optimize serverless data workflows, orchestrating complex ETL processes** across AWS services such as Glue, EMR, and Redshift, ensuring scalability and reliability
- Implemented a **CI/CD pipeline** using AWS CodePipeline and CodeBuild, automating the deployment process and enforcing code quality standards
- Collaborated with cross-functional teams to align data engineering solutions with business objectives and product roadmap
- Technical stack: **Terraform, AWS (Glue, Redshift, Lambda, API Gateway, CloudFormation), Apache Airflow, Python, SQL/NoSQL**

Althea Consulting

Paris, France

MACHINE LEARNING ENGINEER - FULL STACK (FREELANCE)

March 2023 – June 2023

- Realization of projects for various customers with a full stack vision (end-to-end ML projects)
- Solution for reading documents in handwritten Arabic (birth / death documents), and development of a pipeline for parallelizing the process of reading these documents (**OCR (LSTM model) + Langchain + AWS Deployment**)
- Website project to find the most similar celebrity (CNN pre-trained + Cosin Similarity for retrieval), and perform morphing between two faces (**StyleGAN3 pre-trained on FFHQ Dataset**), Deployment (**Django+ AWS**)
- E-commerce platform for AI-generated images and NFTs (**Django, AWS**), enabling the purchase and resale of these items, blogging and the creation of new and unique images generated by GANs
- Stack Technique: **Keras, Git, Docker, Django, AWS cloud (RDS,S3,Sagemaker)**

D3S : Data Science Strategic Services

Grenoble, France

DATA SCIENTIST

August 2022 – March 2023

- Led the formulation of business requirements for a large-scale project automation initiative (Digital Costing), bridging customer needs with ML solutions
- Designed and implemented state-of-the-art architectures for diverse ML problems, including **Time Series Analysis, Computer Vision** (Object Detection), and **Natural Language Processing** (OCR)
- Engineered end-to-end ML pipelines using **AWS SageMaker**, encompassing data preprocessing, model training, and deployment for production environments
- Developed and optimized data ingestion and processing workflows using **AWS Glue, AWS Lambda, and Amazon S3** for efficient big data handling
- Implemented CI/CD practices for ML projects using **Git, Docker, and AWS CodePipeline**, ensuring seamless integration and deployment
- Technical stack: **PyTorch, AWS SageMaker, Git, Docker, OCR, Weights & Biases, AWS Glue, Amazon S3, AWS Lambda**

Two-i : Video Surveillance

Metz, France

MACHINE LEARNING ENGINEER

April 2021 – September 2021

- Implemented advanced Deep Learning architectures for video anomaly detection, including **Encoder/Decoder** and **Convolutional LSTM** models
- Developed production-ready ML pipelines for client projects, focusing on **Object Detection** using **YOLOv4**
- Utilized **AWS SageMaker** for model training and deployment, integrating with **Amazon S3** for data storage and **AWS Lambda** for serverless processing
- Technical stack: **Keras, TensorFlow, AWS SageMaker, Git, Docker, Amazon S3, AWS Lambda**

LIP6 <https://github.com/SaadTazroute/JurisprudencesClassification>

Paris, France

DATA SCIENCE INTERN

April 2020 – September 2020

- Exploration of state-of-the-art techniques for Natural Language Processing (NLP). (**Word Embedding , TF-IDF**)
- Use of Machine Learning algorithms and neural networks for text classification in the juridical field. (**BERT, CamemBERT, Scikit-learn**)
- Technical stack : **Keras, Transformers**

Skills

Languages Arabic (Native), French (Advanced), English (Advanced)(Toeic 950/990)

IA Computer Vision, Natural language processing, Time Series, Generative AI, Transfer learning, Attention Mechanisms

Python Back-end (**Django**), Scientific libraries (**Numpy,Pandas,Scipy**) Deep Learning packages (**PyTorch, Tensorflow**)

AWS **GLUE, Lambda/Step Functions, Sagemaker, RDS/Dynamo DB, AWS CodePipeline (CI/CD)**