Ihab Safouat

Casablanca, Morocco · GitHub · LinkedIn

**MLOps Engineer | Data Scientist | AI Pipeline Specialist**

Building scalable, automated, and robust machine learning systems.

# Technical Skills

**ML Frameworks & Languages:** PyTorch, TensorFlow, Python **MLOps Tools:**

Docker, Kubernetes, MLflow, CI/CD, Airflow, GitOps

## Cloud Platforms:

AWS, GCP, Azure

## Data Processing:

Pandas, Spark, SQL, ETL pipelines

## Other:

REST APIs, Linux, Git, Terraform

# Professional Experience

## MLOps Engineering Intern

*Casablanca, Morocco*

* Developed automated ML pipelines with CI/CD for model training and deployment.
* Managed Dockerized model serving in Kubernetes, with monitoring and rollback.
* Collaborated with data engineers and software teams to integrate production-ready solutions.

## AI Engineer

*Dowslake*

* Developed and deployed a healthcare chatbot using Rasa framework for a multinational hospital network.
* Implemented natural language processing to handle patient inquiries and appointment scheduling.
* Collaborated with healthcare professionals to ensure compliance with medical protocols and data privacy.

# Projects & Research

## LLM Architecture from "Attention is All You Need"

* Implementing transformer from scratch: multi-head self-attention, positional encoding, encoder–decoder.
* Integrated model training flow using modular pipelines, version control, logging, and checkpointing.

## Intrusion Detection System (IDS) with Auto-Retraining

* Built an IDS detecting network anomalies using ML techniques.
* Automated data ingestion, retraining, validation, and deployment using CI/CD workflows.

## Data■Centric AI Pipeline Implementation

* Applied Andrew Ng's "Data-Centric AI" methodology.
* Designed experiments optimizing data quality, labeling, and augmentation.
* Configured monitoring dashboards to track dataset performance drift and pipeline metrics.

## AI-Powered Economic Forecasting System

* Developed a novel hybrid model combining transformer architectures with traditional econometric models.
* Implemented real-time market sentiment analysis using advanced NLP techniques on financial news and social media.
* Achieved 15% higher accuracy in market trend prediction compared to industry benchmarks.

## Distributed MLOps Pipeline for Large-Scale Model Training

* Architected a fault-tolerant distributed training system using Ray and Horovod.
* Implemented dynamic resource allocation and auto-scaling for GPU clusters.
* Reduced training time by 60% while maintaining model accuracy through efficient data parallelism.

## Advanced Game AI with Reinforcement Learning

* Built a sophisticated game AI system using Deep Q-Learning and Monte Carlo Tree Search.
* Implemented multi-agent coordination and adversarial training for complex game scenarios.
* Developed a novel reward shaping mechanism that improved learning efficiency by 40%.

## Multilingual NLP Pipeline for Low-Resource Languages

* Created a transfer learning framework for under-resourced languages using cross-lingual embeddings.
* Implemented zero-shot learning techniques for language understanding tasks.
* Achieved state-of-the-art results in sentiment analysis for 5 low-resource languages.

## Helferlein (NGS Data Workflow Tools)

* + Contributed to the Helferlein repository, automating genome sequencing data processes via custom scripts.
  + Implemented GitHub Actions for CI/CD orchestration of bioinformatics workflows.
  + Enhanced workflow reliability and reproducibility through automated testing and documentation.

# Education & Certifications

## Master's Degree in Big Data Bachelor's Degree in Big Data

* Microsoft Certified: Generative AI with GitHub
* MLOps Certification with GitHub
* LinkedIn Learning Certifications
* Udemy Professional Certifications

# Additional Highlights

* Open source contributions: code and documentation on GitHub for LLM and IDS work.
* Published posts and technical write-ups on LinkedIn to share learnings.
* Strong focus on automation, reliability, and clean infrastructure.