

Anas Hafeed

Machine Learning Engineer | anas.h.hafed@gmail.com | [Ln/LinkedIn](#)

SUMMARY

Machine Learning 4+ years of experience building and deploying production-grade ML systems. I specialize in architecting scalable ETL pipelines from 3,000+ sources and designing NLP systems that integrate over 140 heterogeneous datasets. I have a proven track record of slashing ML training times by 70% through pipeline refactoring, hardware optimization, and distributed training. With deep expertise in RAG architectures—including models like Arabert, Gemini, and Qwen—and a background as a Project Consultant, I bridge the gap between complex algorithmic design and applied AI solutions by mentoring teams to deliver high-impact results

EXPERIENCE

Machine Learning Engineer

Dataplume — Remote — Riyadh, Saudi Arabia

Sep 2022 – Present

- **Architected high-throughput ETL pipelines** to ingest and process data from **3,000+** heterogeneous sources (APIs, web scraping, third-party feeds), ensuring 24/7 data availability for production models.
- **Engineered an Enterprise RAG Framework** integrating **140+** datasets; fine-tuned and deployed LLMs, including **Arabert, Gemini, and Qwen**, to significantly enhance chatbot contextual accuracy and response relevance.
- **Optimized ML Infrastructure**, achieving a **70% reduction in training latency** by implementing distributed training protocols, feature engineering refactors, and strategic hardware resource allocation.
- **Standardized Production Workflows** by implementing robust A/B testing frameworks and real-time model monitoring to ensure the reliability and drift-resistance of deployed AI solutions.
- Collaborated cross-functionally to translate business requirements into scalable ML solutions.

Instructor & Project Consultant

Damascus University

Nov 2023 – May 2025

- Delivered algorithms, software engineering, and data science courses to **150+ students** each semester.
- Supervised ML capstone projects and mentored teams to deliver applied AI solutions.

Freelance Data Scientist / Full-Stack Developer

Oct 2021 – Nov 2023

- Delivered end-to-end ML solutions from data collection to deployment.
- Built ML-powered dashboards and applications with **95%+ client satisfaction**

PROJECTS

Quality Prediction System

- Forecasts PM2.5 levels using geospatial and time-series data (OpenAQ, Open-Meteo).
- Engineering features (lagging, rolling windows, cyclical encoding).
- Compared **XGBoost, LightGBM, LSTM, and SARIMAX**; the ensemble achieved **$R^2 = 0.88$** .
- Built an interactive Streamlit app for real-time forecasting.

Sentiment & Competitive Intelligence Platform

- Scraped multi-region Amazon reviews at scale.
- Applied **FinBERT** and **LDA** for sentiment and topic analysis.

Customer Segmentation & LTV Prediction

- Implemented **RFM segmentation** and **XGBoost** LTV prediction (**$R^2 = 0.85$**)
- Identified 7 customer segments (Champions, At-Risk, etc.).
- Delivered actionable recommendations via Streamlit dashboard

Also applied deep learning to computer vision and NLP, including CNNs, GANs, and language models. Built practical systems using AutoML, predictive modeling, and collaborative filtering for recommendations and customer insights.

SKILLS

Languages & Tools: Python, SQL, Git

ML: Classification, Regression, XGBoost, LightGBM, Feature Engineering

DL & NLP: PyTorch, TensorFlow, CNNs, GANs, Transformers, HuggingFace

MLOps: MLflow, Docker, CI/CD, Model Monitoring, A/B Testing

Data Engineering: ETL, Web Scraping, PostgreSQL, MySQL

Cloud: AWS (S3, EC2, SageMaker)

Large Language Models: Arabert, Gemini, Qwen, HuggingFace Transformers, Prompt Engineering

Languages: Arabic (Native), English (Fluent)

EDUCATION

B.Sc. Applied Science (Computer Technologies)

Minor: Applied Machine Learning — Damascus University

Associate Degree, Applied Computer Technologies

Damascus University — 2022 | Honors (GPA: 4.08)

CERTIFICATION

Machine Learning Specialization — Stanford Online (2023)