AbdulRahman Adham Rashed

Junior AI / Machine Learning Engineer | Data Scientist | NLP & Generative AI

Alexandria, Egypt | abdulrahman.adham17@gmail.com
github.com/Z4iM • linkedin.com/in/abdulrahman-adham

PROFESSIONAL SUMMARY

Information Systems graduate specializing in applied Data Science, Machine Learning, NLP and Generative AI. Built NEXUS EDI, a bilingual multi-agent document intelligence platform integrating hybrid RAG, OCR, secure data pipelines, human review and MLOps. Additional experience in supervised and unsupervised learning, transformer fine-tuning, experimental evaluation and analytics dashboards. Seeking junior AI/ML Engineering, NLP or Data Scientist opportunities.

TECHNICAL SKILLS

Programming & Analytics: Python, R, SQL, Java, C++, Pandas, EDA, preprocessing, feature engineering, ETL concepts

Machine Learning: Scikit-learn, Logistic Regression, Random Forest, SVM, Naive Bayes, K-Means, DBSCAN, cross-validation, ROC-AUC, F1, clustering metrics

Deep Learning & NLP: PyTorch, TensorFlow/Keras, Hugging Face, BERT, DistilBERT, T5, CNN/MobileNetV2, ViT, PEFT/LoRA, TF-IDF, OCR

GenAI & Retrieval: Gemini, LangGraph, function calling, human-in-the-loop, hybrid RAG, BM25, multilingual MiniLM, ONNX, RRF, BGE-M3, FAISS/HNSW, KeyBERT

Backend & Data: FastAPI, Next.js, TypeScript, PostgreSQL, SQLAlchemy, Alembic, MySQL, SQLite, MongoDB, Supabase, REST APIs, RBAC

MLOps & Tools: Docker/Compose, GitHub Actions, Pytest, Playwright, Git, Jupyter/Colab, model evaluation, tracing and monitoring

Visualization: Power BI, Tableau, Matplotlib, Seaborn, R Shiny, ggplot2, QGIS

SELECTED AI, MACHINE LEARNING & DATA SCIENCE PROJECTS

NEXUS EDI — Enterprise Document Intelligence | Flagship AI Engineering Project (2026)

Python · Gemini · LangGraph · Hybrid RAG · NLP/OCR · FastAPI · Next.js · PostgreSQL · Docker

- Built a bilingual (English/Arabic) document intelligence platform for PDF, DOCX, PPTX, TXT and image intake, OCR, document classification and structured extraction of business data.
- Implemented three Gemini-capable agent roles: a standalone investigator plus LangGraph-orchestrated Document and Policy Specialists, with restricted tools, persistent checkpoints and human-in-the-loop review.
- Designed authorization-scoped hybrid RAG using BM25, 384-dimensional multilingual MiniLM embeddings, cosine similarity and RRF; achieved 87.5% Recall@5 and 0.9635 MRR@5 on 40 synthetic bilingual queries.
- Separated AI suggestions from authoritative outcomes through source-verified citations, human-approved policy rules, deterministic evaluation, audit trails and role-based access control.
- Integrated PostgreSQL/Alembic, durable background workers, Docker Compose, GitHub Actions and observability; verified 401 passing backend tests in the latest local showcase checks.

Customer Churn Prediction & Demographic Segmentation

Python · Scikit-learn · Pandas · EDA · Classification · Clustering

- Built a supervised churn pipeline on 7,043 telecom customers with preprocessing, class balancing, EDA and 5-fold cross-validation; Random Forest achieved 76.7% accuracy and 0.842 ROC-AUC (CV F1: 0.634).
- Segmented 10,000 Census Income records using K-Means and DBSCAN; compared elbow/silhouette and Davies-Bouldin/Calinski-Harabasz metrics (DBSCAN silhouette: 0.297).

Transformer Fine-Tuning — PEFT & LoRA Implemented From Scratch

PyTorch · Hugging Face Transformers · DistilBERT · Vision Transformer

- Implemented Prompt Tuning, Prefix Tuning, P-Tuning, BitFit and IA3 as custom PyTorch modules; reduced DistilBERT trainable parameters from 66.9M to 52K (0.08%) using BitFit.
- Built a modular ViT training framework and implemented LoRA manually (ranks 16/32/64); achieved 98.9% validation accuracy on a 10-class image classification task.

SELECTED PROJECTS (CONTINUED)

Bank Customer Attrition Analytics & Dashboard

R · caret · SVM · Naive Bayes · DBSCAN · R Shiny · ggplot2

- Analyzed 10,127 bank records with cleaning, scaling and exploratory visualizations; trained Naive Bayes (87.3% accuracy; 74.9% balanced accuracy), SVM and DBSCAN models.
- Built an interactive R Shiny dashboard for data exploration, classification results and customer segmentation.

NLU Pipeline — Natural Language to Structured JSON

Python · BERT · T5 · Hugging Face Transformers

- Combined a fine-tuned T5 paraphraser, BERT intent classifier and T5 slot extractor on the Schema-Guided Dialogue restaurant domain; achieved 93.3% intent accuracy (0.93 macro-F1) and evaluated the full JSON pipeline on held-out data.

Independent Retrieval-Augmented Generation Pipeline

Python · BGE-M3 · HNSW · KeyBERT

- Built a separate file-agnostic RAG pipeline with structure-aware chunking, KeyBERT tags, dense BGE-M3 embeddings, HNSW search and source-grounded LLM responses.

AI HR Hiring Agent — Agentic Resume Screening

Python · Gemini · Pydantic · SQLite

- Designed parsing, scoring and critic agents for evidence-based resume ranking; used Pydantic output schemas, duplicate detection, persistent SQLite state and human approval before outreach.

ADDITIONAL PROJECTS

Weather Australia BI Dashboards — Tableau and Power BI analysis of 145,460 weather records, including six worksheet visualizations and two dashboards.

OptiSet — MySQL Optimization — Normalized a four-table relational schema and simplified typical retrieval from six joins to two using foreign-key design.

PROFESSIONAL EXPERIENCE & TRAINING

Web Development & Content Security Intern

Jul 2024

Pharco Corporation · Egypt

- Maintained an internal CMS serving 50+ employees and implemented role-based access controls to protect access and data integrity.

Odoo Technical Bootcamp

Aug 2024

Odoo · Remote (8-hour intensive)

- Studied enterprise architecture, relational data modeling and Odoo 16 ORM, including database module structures and business workflows.

EDUCATION

BSc, Information Systems | Graduated September 2026

Arab Academy for Science, Technology & Maritime Transport (AASTMT) · Alexandria, Egypt

Dual Degree Partner Programme: University of Northampton, UK | GPA: 3.1/4.0 (Very Good With Honors)