

Basil Reda

AI Engineer

Cairo, Egypt basilreda112211@gmail.com +201016810656

github.com/BasilReda linkedin.com/in/basil-reda

Professional Summary

AI Engineer specializing in **advanced RAG architectures**, **LLM post-training alignment (DPO, GRPO)**, and **Reinforcement Learning integration**. Adept at designing multi-agent pipelines and high-fidelity retrieval systems utilizing hybrid search and HyDE. Proven track record of building complex ML algorithms from first principles and deploying production-grade conversational APIs using **PyTorch** and **FastAPI**. Driven by a rigorous approach to system architecture, optimizing both model convergence and retrieval performance for complex, real-world AI applications.

Education

AI Trainee – 9-Month Professional Program, Information Technology Institute (ITI)	Oct 2025 – Present
Bachelor’s Degree in Computer Engineering, Mansoura University Faculty of Engineering – Computer Science Department	Oct 2020 – Jul 2025

Projects

FlexiHire | Graduation Project | [GitHub]

- **Spearheaded** a dual-sided intelligent matching recommender system, deployed as a production-grade FastAPI service with low latency via in-memory caching, connecting job seekers and companies through content-based filtering.
- **Formulated a hybrid search strategy** in **Qdrant** backed by a fine-tuned Sentence Transformer, significantly boosting MRR and precision over the baseline.
- **Developed a multi-agent AI pipeline** utilizing CrewAI and GPT-3.5-turbo to automatically extract, validate, and clean skills from raw CVs and job descriptions.

Serenity | [GitHub]

- **Architected a LangGraph-based multi-agent pipeline** exposed via a FastAPI streaming interface, orchestrating complex NLP workflows including 176-language detection, Albert-based emotion classification, and LLM-powered safety guardrails.
- **Engineered an advanced hybrid RAG system** leveraging Pinecone for combined sparse (BM25) and dense vector search, augmented with cross-encoder reranking, relevance grading, and an automated query rewrite loop.
- **Fine-tuned a Qwen2.5 LLM using Direct Preference Optimization (DPO)** to align the model’s generation with safe, empathetic domain-specific guidelines, seamlessly integrating the local checkpoint alongside low-latency Groq API routing.

Acoustic Scene Classification (DCASE 2016) | [GitHub] | [Paper]

- **Modernized legacy architectures to PyTorch**, leveraging **torchaudio** for GPU-accelerated Mel-spectrogram feature extraction to train classification models across multiple environmental sound classes.
- **Established a robust evaluation framework** utilizing **4-fold cross-validation** and automated experiment tracking with **Weights & Biases (W&B)** to monitor loss curves and model convergence.
- **Conducted systematic ablation studies** to optimize inference speed and classification precision, analyzing the impact of **sampling rates** and **He-Kaiming initialization**.

Pac-Man DQN | [GitHub]

- **Authored a Deep Q-Network (DQN) from scratch** using **TensorFlow** and the OpenAI Gym library, orchestrating the full reinforcement learning lifecycle to train an autonomous agent to master Ms. Pac-Man.
- **Designed a custom Convolutional Neural Network (CNN)** for raw pixel state processing, optimizing game frame feature extraction via custom preprocessing pipelines.
- **Implemented critical RL stability mechanisms**, including an **Experience Replay** buffer to break temporal correlations and an **epsilon-greedy decay** strategy to balance exploration and exploitation.

Technical Skills

Core Development: Python, SQL (PostgreSQL, MySQL), NoSQL (MongoDB), Django, FastAPI, Flask, REST APIs, Linux/Bash, Git

Generative AI & Agents: LLMs (GPT-4, Gemini, Llama), LangChain, LangGraph, CrewAI, Haystack, n8n, Hugging Face, RAG Optimization (Ragas, TruLens), PEFT (LoRA/QLoRA)

Machine Learning & RL: PyTorch, TensorFlow, Keras, Scikit-learn, XGBoost, PPO, GRPO, DPO, DQN, Gymnasium (OpenAI Gym), Stable Baselines3 (SB3), TRL

Data Engineering & Vector Data: Pandas, NumPy, Pinecone, Qdrant, ChromaDB, FAISS, Feature Engineering, Data Wrangling

MLOps, Audio & Signal: Docker, MLflow, Weights & Biases (W&B), TorchAudio, Librosa, OpenCV, Prophet, StatsForecast

Analytics & UI Deployment: Tableau, PowerBI, Streamlit, Gradio, Chainlit, Matplotlib, Seaborn, Plotly

Soft Skills

- Leadership
- Communication
- Problem-Solving
- Collaboration
- Adaptability

Language

Arabic: Native | **English:** C1

Additional Information

Military Status **Exempted** (Certificate available upon request).