Saeed Neamtallah

Last updated in December 2025

Giza, Egypt | saeedneamtallahsaeed@gmail.com | 01098095104 | github.com/SaeedNeamtallah
linkedin.com/in/saeed-neamtallah

Summary

AI & **LLM Engineer** focused on shipping **GenAI applications, RAG systems, and high-quality prompt engineering**. Built a production-style RAG pipeline around OpenAI/Hugging Face and local models (Ollama) with **FastAPI**, **Docker**, and **PostgreSQL + pgvector**. Strong deep learning background with **PyTorch** (CNNs, RNNs, NLP, and computer vision) through intensive, project-based internships and advanced diplomas. Comfortable owning the full lifecycle from data preparation and modeling to evaluation and deployment. Target roles: **AI Engineer, LLM Engineer, Prompt Engineer**.

Projects

Mini RAG Pipeline (LLM Application) *FastAPI, Docker, PostgreSQL + pgvector, OpenAI/HF, Ollama*

- Designed and implemented an end-to-end **Retrieval-Augmented Generation** system for question answering over custom documents.
- Engineered and iterated on **prompt templates** for retrieval-aware answering, context injection, and structured response formatting.
- Built a modular pipeline for document ingestion, preprocessing, chunking, embeddings generation, and **pgvector**-based similarity search.
- Exposed clean REST endpoints with **FastAPI** for retrieval and generation, integrating both hosted LLMs (OpenAI/HF) and local models via **Ollama**.
- Added basic evaluation datasets and tracing to compare embedding models and retrieval quality (different embedding backends and distance metrics).
- Containerized the API and database with **Docker & Docker Compose** for reproducible local and server deployments.

Micrograd (from scratch) *NumPy, Autograd/Backprop, PyTorch-style API*

- Implemented a minimal deep learning engine with computational graph construction, **automatic differentiation**, and **reverse-mode backpropagation**.
- Trained simple MLPs via gradient descent, consolidating understanding of optimization, initialization, and numerical stability.
- Strengthened low-level intuition behind frameworks like PyTorch, improving debugging and model-design skills for real-world AI systems.

Experience

Machine Learning Engineer Intern – Digital Egypt Pioneers Initiative (DEPI) Apr 2024 – Oct 2024

- Built and trained deep learning models using **PyTorch**, including CNNs for image classification and RNNs for sequential data.
- Developed end-to-end ML pipelines: data preprocessing, feature engineering, model training, and evaluation.
- Applied **transfer learning** using pre-trained architectures for domain-specific tasks in computer vision.
- Collaborated with cross-functional teams to turn problem statements into deployable ML solutions.

Machine Learning Engineer Intern – Microsoft Student Club (EELU) May 2024 – Sep 2024

- Trained deep learning models with **PyTorch** for classification tasks, experimenting with CNN and RNN architectures.
- Applied data augmentation and hyperparameter tuning to improve model robustness and generalization.
- Wrote modular, production-style code with **Git**-based workflows (branching, PRs, reviews).

Education

B.Sc. in Computer Science & IT **Egyptian E-Learning University (EELU)** Aug 2021 – Oct 2026
Expected Graduation: 2026

Skills & Technologies

LLMs & GenAI: LLM-based applications, RAG systems (chunking, embeddings, vector search with **pgvector**), prompt engineering (context-aware prompts, answer formatting), OpenAI API, Hugging Face models, local LLMs with **Ollama**.

Machine Learning & Deep Learning: PyTorch, NumPy, Pandas, scikit-learn, CNNs, RNNs, transfer learning, model evaluation/validation, feature engineering, hyperparameter tuning.

Backend & MLOps Basics: FastAPI, Docker, Docker Compose, PostgreSQL, database design, Linux, Git, basic CI-style workflows and production-style project structure.

Other: Problem solving, clean code and OOP in Python, technical communication in English (Business English).

Certifications

- **Core AI & ML Diplomas**
  - **Machine Learning Diploma** – CSkilled — Jan 2025 – Oct 2025
    Advanced machine learning with PyTorch: regression, classification, deep learning, and computer vision with focus on data preprocessing, model optimization, and end-to-end ML pipelines.
  - **AI & Data Science – IBM Track** – Digital Egypt Pioneers Initiative (DEPI) — Apr 2024 – Oct 2024
    IBM-style data science methodology: data analysis, machine learning, and predictive modeling using Python and scikit-learn on industry-aligned projects.
- **Deep Learning & NLP Specializations**
  - **Deep Learning for Natural Language Processing** – Udemy
    Word representations, sequence models, seq2seq, attention, transformer networks, and transfer learning for NLP.
  - **Deep Learning for Computer Vision** – Udemy
    CNN-based computer vision: image preprocessing, classification, object detection (YOLO, SSD) and semantic segmentation (U-Net).
  - **Deep Learning MiniCamp [Arabic]** – Udemy
    Practical foundations of neural networks, activation functions, backpropagation, and optimization.
- **Machine Learning Foundations**
  - **Practical Machine Learning for Data Scientists** – Udemy
    Supervised and unsupervised learning (regression, classification, clustering, PCA) with end-to-end workflows and ensemble methods (bagging, boosting).
  - **Python and OOP Diploma** – CSkilled — Jan 2024 – Dec 2024
    Strong Python fundamentals, OOP, modular design, and problem solving through mini-projects.
- **Other Relevant Tracks**
  - **Business English Track** – SYE English Community (DEPI) — Apr 2024 – Oct 2024
    Business-oriented English communication skills for professional and global environments.
  - **AI for Startup** – Udemy
    How to apply AI capabilities in startup and product contexts.

Achievements & Recognition

Top 12 Nationwide – EEG Brain-Computer Interface Competition — Mar 2025

- Ranked in the top **6%** nationally out of 200+ participating teams.
- Engineered signal-processing pipelines for EEG pattern recognition and built deep learning models for brain-computer interface tasks.
- Demonstrated ability to handle complex biomedical AI problems and deliver results under competitive, time-constrained conditions.