Marwan Gamal Hassan

Applied AI Engineer · Deep Learning & Natural Language Processing

Alexandria, Egypt | +20 109 798 2853 | marawangamal229@gmail.com | linkedin.com/in/marwan-gamal-5b40b4280 | github.com/marwan679
ORCID: 0009-0009-7313-2366 | Military Service: Exempted | IELTS Academic 7.5 (C1) | German: Proficient

PROFESSIONAL SUMMARY & OBJECTIVE

Applied AI Engineer with 3+ years of experience across production model deployment, data engineering, computer vision, and NLP. Builds and ships AI systems from ETL pipelines and model training through scalable, production-grade cloud deployment and monitoring — currently leading engineering for an AI-driven cybersecurity platform and a climate-AI analytics initiative. Experienced deploying AI models on edge devices for low-latency, resource-constrained environments. Skilled in model interpretability and explainable AI techniques to support compliance, alongside data governance practices for production-grade data engineering. Three-time published researcher (IEEE, Springer) in deep learning and applied ML.

CORE COMPETENCIES

Applied AI Engineering · Deep Learning & NLP · MLOps & Model Deployment · Cloud Infrastructure (AWS, Azure, GCP) · Data Engineering & ETL · Model Interpretability & Explainable AI · Computer Vision · LLMs & RAG · Team Leadership & Cross-Functional Collaboration

TECHNICAL SKILLS

Programming: Python, R, C, C++, Java, SQL, Git · Design Patterns, SOLID, OOP
Deep Learning: TensorFlow, PyTorch, NAS, Recommendation Systems, GPU Acceleration (CUDA, TensorRT)
LLMs & GenAI: HuggingFace, Ollama, LangChain/Graph, LlamaIndex, RAG, Fine-tuning, MoE, Vector Databases (FAISS, Pinecone, Weaviate)
Computer Vision: YOLO, Detectron, FOMO, MediaPipe, AR/MR/VR
NLP: Gensim, BERTopic, Sentence Transformers, TextBlob, Statistical Language Models
Model Deployment & Serving: TensorFlow Serving, Docker, Kubernetes, Kubeflow, AWS SageMaker, FastAPI (REST APIs), Model Optimization (Quantization, Pruning, TensorRT)
Data Engineering & MLOps: Multimodal Data Pipelines, ETL Pipelines, Data Lake, Feature Store, Model Registry, Data Governance, Model Governance, GitOps, Data Versioning (DVC), Experiment Tracking (MLflow), Monitoring (Prometheus, Grafana), CI/CD, Kafka, TFX, TFLite, GitHub Actions
Cloud Platforms: AWS, Azure, GCP
Methodologies: Agile, Scrum, A/B Testing
Security & AI Hybridization: Autonomous AI agents integrated with cybersecurity tooling

EXPERIENCE

Lead AI Engineer | *EchoAI (Part-Time)* 01/2026 – Present · Alexandria, Egypt

- Founded and led a 5-person team to build Echo Sentinel, an autonomous AI security agent integrating CVE-MCP and threat-graph analysis.
- Led and mentored the engineering team, fostering cross-functional collaboration across 3 functional areas (security, data, and infrastructure) to deliver AI solutions on schedule.
- Deployed Echo Sentinel to production with A/B-tested detection strategies, reducing false positives by 15%.
- Grew Echo Sentinel's detection coverage to 2 tracked CVEs at 98% detection accuracy.
- Architected production-grade infrastructure across 3 cloud platforms (AWS, Azure, GCP) using GitOps-based CI/CD workflows for infrastructure and model versioning; used SageMaker for model training and Kubeflow for pipeline orchestration.
- Used MLflow-based experiment tracking to monitor model performance across 12 training runs per Agile/Scrum sprint cycle.
- Established real-time model monitoring with Prometheus and Grafana, reducing drift incidents by 30%.
- Directed Echo Labs, EchoAI's climate-tech branch, partnering with two external collaborators (NIOF and EU research partners) on climate-AI analytics and oceanographic data engineering.
- Built Echoculus, a long-horizon time-series forecasting framework combining deep learning and agentic systems.

AI Engineer (Contract) | *Extroconnect* 04/2022 – 01/2026 · Remote

- Implemented AI solutions for US real estate, creating multimodal data pipelines that processed 10 TB of data monthly.
- Improved property valuation accuracy by 12% and reduced manual appraisal time by 40%.
- Designed a comp-finding algorithm combining image similarity, textual attributes, and geo-coordinates, achieving 97% match relevancy.
- Applied MLOps best practices (CI/CD, model versioning, experiment tracking), reducing operational costs by 10% while delivering reliable AI services to external organizations.

Embedded Systems Training (Intern) | *ITI & Zatech (Part-Time)* 07/2024 – 09/2024 · Alexandria, Egypt

- Developed C and AVR microcontroller expertise through 210 hours of hands-on training in IoT and RTOS, reducing firmware development time by 30% on subsequent projects.
- Integrated hardware and software components for real-time systems, supporting system-level development and testing.

AI Training | *ST Smart Technology (Part-Time)* 04/2024 – 06/2024 · Alexandria, Egypt

- Developed computer vision skills for edge/IoT devices, autonomous navigation, and sensor fusion over 50 hours, achieving 92% real-time navigation accuracy in coursework.
- Integrated CV models on microcontrollers and control systems.

AI Training | *Zewail City, Valley of Science & Technology (Part-Time)* 08/2023 – 10/2023 · Hybrid, Egypt

- Engineered a CNN from scratch using pure linear algebra, with no ML frameworks, as the capstone project of a 50-hour AI training program.

KEY ACHIEVEMENTS

- Grew Echo Sentinel's production detection coverage to 98% accuracy while cutting false positives by 15% through A/B-tested detection strategies.
- Reduced model drift incidents by 30% by establishing real-time monitoring (Prometheus, Grafana) across production AI systems.

- Improved property valuation accuracy by 12% and cut manual appraisal time by 40% through multimodal data pipelines processing 10 TB of data monthly.
- Cut operational costs by 10% and firmware development time by 30% through applied MLOps practices and hands-on embedded systems training.
- Achieved 97% match relevancy in a comp-finding algorithm combining image similarity, text, and geo-coordinate data.

EDUCATION

BSc Computer Science & AI, Pharos University, Alexandria **2021-2025**
CGPA: 3.6 Grade: Excellent with Honors

MS Computer Science , Alexandria University , Alexandria **2026-present**

KEY PROJECTS

Voice Intelligence Pipeline - Whisper · Voxtral · Spectral Clustering

- Built an end-to-end voice agent supporting speaker diarization, ASR transcription, audio understanding, and function calling directly from voice input.
- Integrated Mistral AI's Voxtral for audio reasoning; applied spectral clustering for robust multi-speaker segmentation on real audio datasets.
- Implemented real-time inference pipelines using TensorFlow Serving and Docker containers, achieving low-latency voice-command processing.

refined-augment - Published Python Package (PyPI)

- Designed and deployed an open-source Python package enabling plug-and-play Augmented Reality integration into computer vision pipelines.
- Engineered the package entirely from scratch using pure Python, linear algebra, and projective geometry, with zero external AR dependencies; installable via pip install refined-augment.

Financial Topic Intelligence - BERTopic · Ollama Llama 3.2 · UMAP · HDBSCAN

- Engineered an NLP data pipeline on the 2026 AAI FinEcon dataset using BERTopic with thenlper/gte-small embeddings for dynamic topic discovery.
- Reduced dimensionality via UMAP and applied HDBSCAN for noise-robust clustering.
- Automated topic labeling via Ollama Llama 3.2 with KeyBERT-inspired reranking on Tesla T4/CUDA.
- Evaluated on a 30-sample test set achieving 86% faithfulness, 81% answer relevancy, and 63% correctness (75% overall).

Federated Learning: Diabetic Retinopathy Detection - Flwr · TensorFlow · TensorRT

- Built a privacy-preserving federated system upholding data privacy by keeping patient data on-device, enabling clinical organizations to train state-of-the-art models without centralizing sensitive data.
- Eliminated data centralization costs using distributed computing across nodes; achieved competitive accuracy via transfer learning, model quantization, and TensorRT optimization.
- Implemented SHAP-based explainable AI dashboards to provide transparent, interpretable predictions for clinical end-users.

PUBLICATIONS

- "**Towards Transparent and Robust Early Diagnosis of Alzheimer's Disease: A Deep Learning and Visual Analytics Framework.**" 2025 Intelligent Methods, Systems, and Applications (IMSA), Jul. 2025. DOI: 10.1109/imsa65733.2025.11167773
- "**A New Era of Artificial Intelligence: Enhancing Video Quality with Cutting-Edge Techniques.**" IEEE ICMISI, May 2025. DOI: 10.1109/ICMISI65108.2025.11115279
- "**Augmented Reality-Enhanced Machine Learning for Automated Recycling Detection.**" Lecture Notes on Data Engineering and Communications Technologies (Springer AISI 2024), vol. 233, pp. 144–151, Jan. 2024. DOI: 10.1007/978-3-031-77299-3_14

SELECTED CERTIFICATIONS

- LangChain MasterClass: Build 15 LLM Apps with Python — Packt (Coursera)
- Large Language Models for Business — DataCamp
- Fine-tuning with Llama 3 — DataCamp
- Building Scalable Agentic Systems — DataCamp
- Fine-tune BERT for Text Classification with TensorFlow — Coursera
- Embedded Systems & IoT — ITI Professional Certificate

AWARDS & RECOGNITION

- Won second place at the Startup Olympiad by the ISF Foundation. (2023)
- Presented "Heart Disease AI Prediction" to Egyptian Prime Minister Dr. Mostafa Madbouly at Silicon Waha (2024)
- Won first place at Intra Africa competition for startups.(2024)

PROFESSIONAL AFFILIATIONS

- Member, IEEE

VOLUNTEER EXPERIENCE

- Led the Academic Committee for Hult Prize, Pharos University (2024–2025), closing out the season's programming initiatives.
- Represented Sprints as Student Ambassador (10/2023–12/2025), spreading awareness across partner universities and staffing the Rise Up Summit Day booth.
- Directed the AI Committee for GDSC, Google Developer Student Clubs (2023–2024), leading technical programming for the academic year.
- Delivered a workshop on writing and implementing scientific papers as a Member of the IEEE PUA Student Branch (2025–Present).