Mohammed Sedeg

AI ENGINEER & RESEARCHER

Turkey (open to relocate) | Portfolio | GitHub | Email | Linkedin | Courses & Certifications

AI Engineer & Researcher delivering **production-grade AI systems** from **research to deployment**, with expertise in **LLMs, computer vision, robotics, and edge AI**. Skilled in **model optimization, distributed training, and generative AI.**

Experience

VisionCore | AI Engineer, Computer Vision Specialist **2024 – Current**

- Built low-latency TFLite/ONNX models for Jetson & Pi (**60–80 ms inference**).
- Optimized models (**−40–60% size**, **2× speed**) via quantization, pruning, clustering.
- Automated **25%** of labeling using Roboflow + VLM-based augmentation.
- Developed YOLOv8 PPE detector (**85–90% mAP**) + anomaly detection (RF), **−15%** over-fitting.

Freelancer | Generative AI & LLM Engineer **2023 – Current**

- Fine-tuned GPT-2 & LLaMA (↑**15% accuracy**, ↓**70% cost**) via Hugging Face, LoRA, Unsloth.
- Built scalable RAG stacks (Qdrant, FastAPI, MongoDB) for context-rich generation.
- Deployed quantized LLMs (gguf, ExLlamaV2) for **~3× faster edge inference**.
- Improved LLM outputs by **20%** using CoT & ToT prompting.
- Cut training time **20%** via PyTorch DDP, AMP, optimized loaders.

Key Projects

- MyLLM–Modular PyTorch LLM framework: training, fine-tuning, RLHF, quantization.
- SilvaXNet– Lightweight DL framework for GPU/CPU educational demos (Python, NumPy, CuPy)
- RagApp – Enterprise RAG system: multi-provider, scalable deployment (Python, FastAPI, Docker, Qdrant, MongoDB, Postgres, SQLAlchemy, Alembic).

Education

M.Sc. Mechatronics – Computer Vision | Karabük University – 2023

B.Sc. Electrical Engineering – Control Systems | Sudan University – 2016

Skills & abilities

- **Programming:** Python, C, C++, MATLAB.
- **ML/DL Frameworks:** PyTorch, TensorFlow, JAX, Scikit-learn.
- **AI & NLP:** Hugging Face, LangChain, LoRA, PEFT, RAG, FAISS, OpenAI, Cohere
- **Computer Vision:** YOLOv8, OpenCV, Ultralytics, Roboflow, Detectron2, Albumentations.
- **Data & Workflow:** PostgreSQL, SQLAlchemy, Alembic, PySpark, Pandas, MongoDB, MLflow.
- **MLOps & Deployment:** Docker, FastAPI, CI/CD, GitHub Actions, AWS SageMaker.