Phu Tran, PhD

Scottsdale, AZ | tr.ngocphu@gmail.com | (818) 927-9957 | tranngocphu.github.io | tranngocphu
phungoctran | Google Scholar

Summary

AI Engineer and Data Scientist with deep experience designing, building, and deploying end-to-end machine learning and agentic AI systems. Skilled in LLM integration, reinforcement learning, and scalable cloud-native AI solutions. Proven ability to translate complex problems into deployable solutions, build robust MLOps and CI/CD pipelines, and collaborate effectively across interdisciplinary teams. **No visa sponsorship required**.

Professional Experience

AI Engineer / Research Associate, Mayo Clinic Arizona – Scottsdale, AZ, USA — June 2025 – present

- Design, build, and deploy agentic AI systems that retrieve and reason over patient EHR data (FHIR, HL7) across clinical workflows.
- Develop AI-powered clinical trial matching and automated cancer registry enrollment solutions.
- Architect scalable full-stack AI platforms using Azure OpenAI, Google Gemini, Python (FastAPI), Next.js, Docker, and Google Cloud (Vertex AI, Cloud Run, Cloud Build, Cloud SQL, BigQuery).
- Implement LangChain- and LangGraph-based multi-agent pipelines for clinical decision support.
- Published highly rated open-source private LLMs integration in LangChain & LangGraph.

Postdoctoral Associate, Brandeis University – Waltham, MA, USA — Feb 2022 – Apr 2025

- Led AI research projects developing ML models for forecasting and control of bio-inspired materials.
- Built CNN/RNN models for video-based velocity estimation, outperforming rule-based systems by large margins.
- Designed Transformer and quantized CNN autoencoder architectures to predict long-term material dynamics with 2× greater horizon accuracy.
- Developed reinforcement learning (PPO, CNN-based policies) for closed-loop control of active materials in experiments and simulations.

Research Fellow, Nanyang Technological University – Singapore — May 2018 – Jan 2022

- Led ML projects for real-time aircraft trajectory prediction and AI-assisted air traffic control systems.
- Achieved 30% accuracy improvement via data augmentation and enhanced GRU-based sequence models.
- Built RL-based Human-AI control prototypes showcased at the Singapore Airshow 2020.
- Deployed large-scale MySQL infrastructure (1.6B+ rows) supporting data-driven aviation research.

Skills

- **AI/ML**: Generative AI, LLMs, RAG, Transformers, Reinforcement Learning, Computer Vision
- **Frameworks**: Python, PyTorch, Scikit-learn, LangChain, LangGraph, FastAPI
- **MLOps & Cloud**: GCP (Vertex AI, Cloud Run, BigQuery), Azure OpenAI, AWS (SageMaker), Docker, CI/CD
- **Tools**: MLflow, Weights & Biases, Ray, SQL, Git, DVC, Linux
- **Certificate**: Google Data Analytics Professional Certificate

Services and Other Achivements

- Publication Chair of The 1st International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT 2020)
- Published 20+ peer-reviewed research articles, with 460+ citations on Google Research Scholar

Education

Nanyang Technological University, Singapore, PhD in Mechanical Engineering — Aug 2012 – July 2017

University of Technology, HCMC, Vietnam, BEng in Mechanical Engineering — Aug 2007 – Apr 2012