Dustin Nguyen

dnguyen.phys@gmail.com | dustindnguyen.github.io | linkedin.com/dustinnguyen24 | Active Secret Clearance

Summary

Applied ML Scientist with Physics PhD and 2+ years building and deploying production ML systems across defense and research. Specializes in surrogate modeling, time-series ML, and scalable PyTorch systems.

Technical Skills

Languages & Frameworks: Python, PyTorch, SQL, Ray, Transformers, scikit-learn, numpy, pandas
GenAI Systems: Retrieval Augmented Generation (RAG), Vector Search, Prompt Engineering, Model Evals
MLOps & Engineering: Git, Docker, Mlflow, Unit Testing
Cloud & Data: AWS EC2 and S3, Data pipelines, Parallel & Distributed Computing (Ray, Lightning DDP)
Research: Time-series models, Transformers, SciML (Neural ODEs and variants)

Experience

Senior Machine Learning Research Scientist, Leidos – Remote, USA   Mar 2025 – Present

- Architected Transformer time-series surrogate model for DARPA enabling real-time Space Domain Awareness at constellation scale - **1000x faster and 30% relative reduction in error** than legacy physics-based methods.
- Designed **sim-to-real training** pipeline and novel feature engineering to ensure generalization under limited real-world data; owned end-to-end ML workflow using MLflow.
- Contributed to development of graph-based feature extraction pipeline to detect coordinated behavior among resident space objects (RSOs) in space constellations.
- Developed internal Python packages with automated CI/CD pipelines and private PyPI distribution for an IARPA remote-sensing program; improved logistic regression model performance for chemical aerosol detection and resolved production bottlenecks.

Senior Machine Learning Engineer, Lockheed Martin – Denver, CO   Mar 2024 – Mar 2025

- Led Python model development for a DARPA project deploying Controlled Neural ODEs as surrogate models in AFSIM simulations. Scaled distributed memory-efficient adjoint-based training with PyTorch Lightning across multi-GPU systems and integrated models into AFSIM using LibTorch. Delivered quantitative results and analyses to government stakeholders through monthly SharePoint presentations.
- Developed multi-modal AI/ML method for a classification task utilizing Meta's data2vec2.0 pre-trained model.

Post-doctoral Researcher, Ohio State University – Columbus, OH   Dec 2023 – Mar 2024

Applied Machine Learning Fellow, Los Alamos National Laboratory – Los Alamos, NM   May 2022 – Aug 2022

- Researched and developed Neural ODEs as a surrogate model for non-linear physics using Julia language.

Graduate Researcher & NASA FINESST Fellow, Ohio State University – Columbus, OH   Aug 2020 – Dec 2023

- Studied idealized galactic feedback models, which is relevant to understanding galaxy evolution, using theoretical and computational methods. Total of 10 published papers, with 6 being first author.
- Awarded competitive NASA FINESST Fellowship (acceptance rate $\sim$8%) for proposed research

Projects

Universal Differential Equations for learning non-linear physics

- Developed the first use case of Neural ODEs (UDEs) within idealized galactic feedback models using Julia.
- Published proof-of-concept model into ICML 2023 Workshop on ML for Astrophysics.
- Applied to real-world data and published into NeurIPS 2023 Workshop on ML and the Physical Sciences.

Courses

LearnSQL.com
- Learn SQL From A to Z Track: 7 comprehensive courses (in prog. ↻).

DeepLearning.ai
- Finetuning & Reinforcement Learning (RL) for LLMs: SFT, RLHF, GRPO, Constitutional AI, etc. (in prog. ↻)
- Retrieval Augmented Generation (RAG): Retriever, Vector Databases, Rerankers, Chunking, etc. (2025 ↗)
- Agentic AI: Multi-agent Workflows, Task Decomposition, Tools, Reflection, Evaluations, etc. (2025 ↗)
- Machine Learning Specialization: ML Fundamentals (2023 ↗)

Education

Ohio State University – PhD in Physics   Dec 2023
Ohio State University – MS in Physics   May 2021
Arizona State University – BSs in Physics and Astrophysics   May 2018