Dustin Nguyen

Physics Ph.D. Candidate with Machine Learning Experience

+1-(480)-823-8799 • dnguyen.phys@gmail.com • in • • • U.S. Citizen

Languages: Python, Julia, JAX, TensorFlow, PyTorch, sci-kit learn, SQL

Expertise: Deep Learning, Natural Language Processing, Numerical Methods, Computational Physics, Data Science

Education

Ph.D. in Physics, The Ohio State University — **Columbus, OH**
08/2018 - 12/2023

M.S. in Physics, The Ohio State University — **Columbus, OH**
08/2018 - 05/2021

B.S. in Physics and Astrophysics, Arizona State University — **Tempe, Arizona**
08/2014 - 05/2018

Work Experience

Los Alamos National Laboratory, Los Alamos, NM *05/2022 - 08/2022*

- As an Applied Machine Learning Fellow, I conducted a study on how neural networks embedded in time-dependent PDEs learn; Identified difficulties in learning unknown physics within chaotic systems.

Projects

Model Discovery with Neural Ordinary and Partial Differential Equations *08/2022 - Present*

- Independent project lead for the first study of neural coupled ODEs that contain singularities; Feature-engineered a custom loss function that additionally penalizes diverging solutions; showed both ADAM and BFGS optimization is required for convergence. Resulted in ICML 2023 workshop paper.
- Application of above models to real observations of galaxies. Intend on submitting to NeurIPS Physical Sciences.

Large Language Model (LLM) and Generative AI projects | *07/2023 - Present*

- Used Cohere API for various projects including sentiment analysis and semantic search.
- Toxicity reduction fine-tuning of FLAN-T5 LLM with PEFT and RLHF on SageMaker AWS (DeepLearning.ai).

Protein Function Prediction (kaggle: CAFA 5) | *05/2023 - Present*

- Developed a model to predict protein functions; tested Ridge, Decision Tree, and Multi-layer Perceptron regression models on T5, ESM2, and ProtBERT embeddings; submitted to CAFA 5 kaggle code competition.

Three-Dimensional Hydrodynamic Simulations of Feedback from Galaxies (PhD thesis) *08/2019 - Present*

- Project lead on computational studies on understanding the physics of driving large-scale galactic superwinds and phenomenological implications of different models using Python and C++. Resulted in 4 first-author papers.
- Used GPU accelerated code `Cholla` to run 3D time-dependent hydrodynamic simulations on supercomputers.

Publications

Machine Learning (1 Independent Author)

- "Neural Astrophysical Wind Models," Nguyen, 2023, ICML 2023 Workshop on M.L. for Astrophysics.

Astrophysics (4 First Author, 2 Co-author, 2 in prep.)

- Nguyen et al. 2023 (arxiv), Nguyen et al. 2023 (MNRAS Letters), Nguyen & Thompson 2022 (The Astrophysical Journal Letters), Nguyen & Thompson 2021 (MNRAS), Co-author on two ApJ papers (2020 and 2023).

Award

- NASA FINESST Fellowship | Student led proposal ∼8% acceptance rate, ∼ $97K — 2022

Machine Learning Training

- **DeepLearning.ai (2023)**: Machine Learning Specialization (3 months), Deep Learning Specialization (in prog.,5 months), Generative AI with LLMs (3 weeks), ML Engineering for Production (MLOps) Specialization (in prog.,4 months), Generative Adversial Networks (GANs) Specialization (in prog., 3 months) | | **Weights & Biases (2023)**: W&B 101. Certificates on in.
- *NLP/LLM Courses (2023):* Free courses offered by Cohere.
- The Erdos Institute Data Science Boot Camp (May 8 2023 - June 7 2023), Team Project: Cafa 5 Protein Prediction.