Fernando Torales Acosta

ftoralesacosta | fernando-torales-acosta | ftoralesacosta.io | fernandotta@gmail.com

Summary

Machine learning research engineer specializing in AI for sciences. Experienced in fast generative diffusion models, AI-assisted design, and deep learning for denoising data. Passionate about using ML to build novel solutions. 7 YOE C/C++, 6 YOE Python, Tensorflow, and PyTorch. Based in NYC.

Work Experience

Lawrence Berkeley National Lab – Machine Learning Research Fellow Feb 2022 - present

- Designed an optimized diffusion model with 150x lower storage requirements and 3x faster sampling
- Contributor to early foundation model in collider physics; transformer model trained on 256 A100s
- Led the construction of synthetic data generation pipeline using C, HDF5, TensorFlow, and Docker
- Developed an ML ensembling module for testing model robustness, reducing model variance by 55%
- Developed a data processing algorithm using GNNs, saving $1M over traditional hardware alternatives

University of California, Berkeley – Graduate Researcher Aug 2018 - Dec 2021

- Developed first Stable Matching Algorithm alternative to k-nearest neighbors for petabyte-scale data
- Implemented parallelized data correction framework to decrease processing from 10 days to 4 hours
- Converted ELT pipeline to ETL for Petabytes of data, reducing team's load times by 4x
- Designed deep learning model for background rejection, beating a robust 15-year-old algorithm by 30%

Projects

OmniFold PyPi: Deep learning for Denoising Data Python Package

Particle Feature Regression with Graph Neural Networks GitHub Link

- Utilized DeepSets and Graph Neural Networks to optimize physical detectors; AI Codesign

Comparing Denoising Diffusion Models GitHub Link

- Trained and compared diffusion models based on CNNs to models based on Point Clouds + GNNs
- Helped change inefficient paradigm of image-based modeling, and influenced later generative models

Education

University of California, Berkeley	Ph.D. (Physics)	2016 - 2021
Stony Brook University	B.S. (Physics)	2012 - 2016

Relevant Publications

- Point Cloud Diffusion for complete event generation: **NeurIPS 2024 #131**
- Point Cloud vs. Image based generative diffusion: **NeurIPS 2023 #192**
- Optimal design of detectors using DeepSets and GNNs: **DOI: 10.1088/1748-0221/19/06/P06002**
- Isolated Photon-Hadron Correlation in ALICE: **Phys. Rev. C 102, 044908**

Skills

Python; C; Git; Pytorch; Tensorflow; scikit-learn; Lua; Docker; Kubernetes; HDF5
Model Evaluation; GPU Optimization; Continuous Integration; Version Control