

# MAURICIO HUMBERTO FERRATO

mferrato@udel.edu | [Website](#) | [Github](#) | [LinkedIn Profile](#)

## EDUCATION

### **University of Delaware**

*Ph.D., Computer and Information Sciences*  
Expected Graduation Spring/Summer 2023

Advisor: [Prof. Sunita Chandrasekaran](#)

[Computational and Research Programming Lab \(CRPL\)](#)

Research Interests:

- Building machine learning (ML) and deep learning (DL) models to study and predict disease outcomes for rare diseases
- Scaling ML and DL models for multi-GPU training on large computer systems
- Using directive-based parallel programming models to accelerate scientific applications

### **University of Delaware**

*M.S., Computer and Information Sciences*

Advisor: Prof. Sunita Chandrasekaran

Newark, DE

September 2018 - Present

Newark, DE

May 2021

### **University of Delaware**

*B.S., Computer Engineering*

Newark, DE

May 2018

## SKILLS

- Computer Languages: Experience with C/C++, Python, Fortran, Java, Julia, R
- Machine Learning/Deep Learning: Scikit-learn, Keras, TensorFlow, PyTorch, Pandas, Matplotlib, Numba, Horovod
- High Performance Computing (HPC): OpenACC, OpenMP, CUDA, Cupy, MPI, HIP, Dask, Profiler Tools
- OS: Windows, Linux, Mac OS
- Other Skills: SLURM, IBM JSM, Anaconda, git, LaTeX
- Languages: Fluent in English and Spanish

## ACADEMIC EXPERIENCE

### **Research Assistant, University of Delaware**

*Position under Dr. Sunita Chandrasekaran*

Newark, DE

September 2018 - Present

- In collaboration with the [Nemours Children's Hospital](#), under Dr. Erin L. Crowgey
- Conducting research on a big data analytics and machine learning project on topics:
  - Developing ML models to predict disease onsets or outcomes for rare type diseases for pediatric patients (acute myeloid leukemia and sickle cell disease)
  - Developing ML models to accurately classify individuals as responders/non-responders to certain drugs prior to treatment (acute myeloid leukemia) (DOI: [10.1093/bioadv/vbad034](#) )
  - Leveraging preprocessing and feature selection techniques to identify biomarkers for disease onset/outcome prediction
  - Exploring advanced ML and DL classification techniques (Random Forests, Gradient Boosting, LASSO, DNNs) that are best suited for complex, highly dimensional genomic data
  - Analyzing the model predictive performance using sensitivity scores, specificity scores, and ROC curves
  - Integrating multi-omics and clinical record data to leverage genetic and environmental factor information that can improve prediction performance
  - Scaling ML and DL models to accommodate for the growing number of large healthcare datasets and data driven healthcare projects
- Oversee undergraduate research projects, providing research expertise in the field of machine learning and data science

### **Standard Performance Evaluation Corporation (SPEC) High Performance Group (HPG)**

*Project Leader, NICAM-DC*

September 2018 – March 2019

- Responsible for integrating the NICAM-DC benchmark into the SPEC harness
- Helped assemble the workload configuration used to run NICAM-DC, a high-resolution global atmospheric model simulator that is a dynamical core package part of NICAM
- Built documentation associated with the benchmark, comprising of compilation/runtime flags, dependencies, and restrictions
- Analyzed runtime of the different workload sizes using multiple different compilers (Intel, Cray, GNU, Nvidia, LLVM) and computer architectures

## Vertically Integrated Projects (VIP), High Performance Computing (HPC), University of Delaware

Newark, DE

Advisor: Dr. Sunita Chandrasekaran

January 2017 – May 2018

- In collaboration with [Department of Chemistry & Biology's Prof. Juan Perilla](#)
- Parallelized a chemical shift prediction algorithm using a portable programming model, OpenACC (DOI: [10.1371/journal.pcbi.1007877](https://doi.org/10.1371/journal.pcbi.1007877))
- Parallelized NASA's NAS NPB Benchmark Suite by implementing a boss-worker model using MPI and OpenACC, allowing for simultaneous computation with multiple GPUs
- Profiled and analyzed runtime of the most time-consuming functions for all projects using PGPROF
- Rewrote, optimized, and redesigned code to account for better memory management
- Created written reports, designed poster presentations, and presented the research projects to a variety of peers and faculty

## WORK EXPERIENCE

### **NVIDIA Corporation**

Santa Clara, CA

Energy Solutions Architect Intern

June 2022 – September 2022

- Worked on a project accelerating the annual energy production (AEP) calculation of wind farm simulations using GPUs.
  - Learned the role of a solutions architect:
    - Used critical thinking skills to find solutions and adapt projects based on the costumers' needs
    - Worked in a team of four, cooperating with business managers and operators
    - Presented new features and advances in the project to partners and costumers regularly
    - Developed proof-of-concept designs to explain and validate the benefit of the work to consumers
  - Accelerated a Python-based wind farm simulation tool (PyWake) to run on GPUs using CuPy
  - Utilized profiler tools such as NVIDIA Nsight Systems and Nsight Compute, and Python's CProfile, to identify the bottlenecks of the code
  - Tested the GPU implementation on state-of-the-art GPUs (V100s, A100s) and showed more than 2x speed up on the multicore 12 CPU core version of the simulation over 200 wind turbines and 4000 wind speed and wind directions
  - Implemented a genetic algorithm optimization approach to prove that many AEP calculations can be ran at the same time on a single GPU, thus providing potentially higher speed up when the AEP calculation is used in an optimization problem

### **Genome Profiling, LLC (GenPro)**

Wayne, PA

Intern

March 2021 – March 2022

- Conducted research on a machine learning project on topics:
  - Developing ML models to accurately classify individuals based on phenotypes for Parkinson's Disease (PD)
  - Leveraging DASK to process a large dataset of 5 million single-nucleotide profiles (SNP) and perform distributed parallelization of model training and prediction
  - Applying the techniques learned through the project collaboration with Nemours on this project to examine model portability across diseases
  - Linking together the multi-omics dataset to methylation scores and designing a mixed data model to outperform the prediction performance of the PD classification model

### **National Cancer Institute (NCI)**

Rockville, MD

Summer Intern

June 2018 – August 2018

- Worked on various parts of Pilot 1 of the [ECP-CANDLE](#) project, including:
  - Estimating model uncertainty using dropout layers at inference as a Bayesian approximation
  - Showing scalability of deep learning model training running on multi-GPU/multi-node systems using Horovod
- Enhanced, compiled, and ran the projects on the NIH's supercomputer Biowulf
- Documented and wrote tutorials for approaching data parallelism and uncertainty estimation in deep learning models based on the project experience

## PUBLICATIONS

1. **Mauricio H Ferrato**, Adam G Marsh, Karl R Franke, Benjamin J Huang, E Anders Kolb, Deborah DeRyckere, Douglas K Graham, Sunita Chandrasekaran, Erin L Crowgey, Machine learning classifier approaches for predicting response to RTK-Type-III Inhibitors demonstrates high accuracy using transcriptomic signatures and ex vivo data, *Bioinformatics Advances*, 2023; vbad034, <https://doi.org/10.1093/bioadv/vbad034>

- Eric Wright, **Mauricio H. Ferrato**, Alex J Bryer, Searles Robert, Juan R Perilla, Sunita Chandrasekaran, (2020) "Accelerating prediction of chemical shift of protein structures on GPUs: Using OpenACC." PLOS Computational Biology 16(5): e1007877. <https://doi.org/10.1371/journal.pcbi.1007877>

#### MENTORING EXPERIENCE

- **Oak Ridge National Laboratory (ORNL) GPU Hackathon - GPU Hackathon Mentor** Knoxville, TN – 2018, 2019
- **Princeton University GPU Hackathon – GPU Hackathon Mentor** Princeton, NJ - 2019
- **Massachusetts Institute of Technology (MIT) GPU Hackathon – GPU Hackathon Mentor** Cambridge, MA – 2019
- **Brookhaven National Laboratory (BNL) GPU Hackathon – GPU Hackathon Mentor** Upton, NY - 2018
- **University of Colorado Boulder (UCB) GPU Hackathon – GPU Hackathon Mentor** Boulder, CO - 2018

#### POSTERS & INVITED TALKS

- **[Invited Talk] Mauricio H. Ferrato**, "How to Approach Your First Data Science Project", Data Science Community Hour, University of Delaware, 2021
- **[Poster] Mauricio H. Ferrato**, Erin L. Crowgey, Adam Marsh, Karl Franke, Sunita Chandrasekaran, "Machine Learning Techniques for Classification of Patient Cohorts Using a Genomics Dataset", DSI Darwin 2021 Symposium, University of Delaware
- **[Poster] Mauricio H. Ferrato**, Erin L. Crowgey, Sunita Chandrasekaran, "Proposing a Machine Learning Framework for Classification of Patient Cohorts Using Genomics Data", AMIA 2020 Virtual Annual Symposium
- **[Poster] Eric Wright, Mauricio H. Ferrato**, Alex Bryer, Robert Searles, Juan Perilla, Sunita Chandrasekaran, "Accelerating Chemical Shift Prediction for Large-scale Biomolecular Modeling", International Supercomputing Conference (ISC) High Performance 2019, Frankfurt, Germany, Best Poster
- **[Poster] Eric Wright, Mauricio H. Ferrato**, Sunita Chandrasekaran, "Accelerating Chemical Shift Prediction for Large-scale Biomolecular Modeling", GPU Technology Conference (GTC) 2019
- **[Poster] Eric Wright, Mauricio H. Ferrato**, Alex J. Bryer, Robert Searles, Juan R. Perilla, Sunita Chandrasekaran, "Estimating Molecular Dynamics Chemical Shift with GPUs", SuperComputing 2018, Dallas, TX
- **[Invited Talk] Eric Wright, Mauricio H. Ferrato**, Thomas Huber, Edwin Ortiz, "GPU Acceleration of Chemical Shift Prediction", Special Interest Group in Software Systems (SIG-SYS) Seminar Talk, University of Delaware, 2018

#### PROFESSIONAL VOLUNTEERISM

- *Student Volunteer*, SC19, Denver, CO
- *Student Volunteer*, SC18, Dallas, TX

#### AWARDS AND HONORS

- **Best Research Poster Award**: Accelerating Chemical Shift Prediction for Large-scale Biomolecular Modeling. Eric Wright, Mauricio H. Ferrato, Alexander J. Bryer, Robert Searles, Juan R. Perilla, Sunita Chandrasekaran. International Supercomputing Conference (ISC) High Performance, June 2019, Frankfurt, Germany
- **Best Research Poster Award**: Vertically Integrated Project (VIP) midatlantic VIP competition. Accelerating Chemical Shift Prediction of Protein Structures using GPUs. Eric Wright, Mauricio Ferrato, Alexander Bryer, Robert Searles, Juan Perilla, Sunita Chandrasekaran. March 2018, Delaware, USA