

# JAMIL GAFUR

Jamilgafur.com ◇ Jamil-gafur@uiowa.edu ◇

Iowa City, IA, 52246

U.S Citizen

## RESEARCH INTERESTS

---

Previously my work has been focused on High Performance and Scientific Computing. I am now transitioning towards the development and optimization of Neural Networks, as well as research in explainable ML using Generative Adversarial attacks and meta-heuristic algorithms.

My interest in these fields came about by observing the issues associated with current Machine Learning (ML) techniques. Currently, there are no standards for developing Neural Models. These models might not be extracting as much information from the dataset or learning incorrect features. Furthermore, the distribution of data is not always equal, and this can cause the models to have biased. Moving forward I am interested in uncovering what is behind the “black box” design of neural networks and using this knowledge on other projects.

## EDUCATION

---

### University of Iowa, Iowa City, Iowa

*Fall 2020 - Current*

- *Ph.D. in Computer Science*

- GPA: 3.48/4.0 <sup>1</sup>

- Featured Classes:

- \* Statistical Machine Learning

- \* Design and Implementation of Algorithms

- \* Scientific Computing and Machine Learning

- \* Optimization Techniques

### CUNY: Lehman College, Bronx, NY

*Fall 2016 - Spring 2018*

- *BS in Computer Science Minor in Business Administration*

- Graduated Cum Laud / Departmental Honors

- GPA: 3.6/4.0

### SUNY: Dutchess Community College, Poughkeepsie, NY

*Fall 2014 - Spring 2016*

- *AS in Computer Science*

- GPA: 3.08/4.0

## AWARDS

---

### GEM Employer Fellowship

*Awarded Fall 2022*

### U-Iowa Computer Science Grant

*Awarded Fall 2021*

## TECHNICAL SKILLS

---

- **Programming Languages**

- Python
  - Bash/Shell
  - Java
  - Fortran 90
  - CUDA

- **Skills**

- GIT
  - Tensorflow/Keras
  - Deep Learning
  - Data Visualization
  - Graph Processing

- **Techniques**

- Neural Network Pruning/Optimization
  - Parallel workflow processing
  - Adversarial Network Development
  - Graph Convolutional Design
  - Heuristic Optimization

---

<sup>1</sup>Due to COVID and my fathers death my second semesters GPA was lower than expected.

## RESEARCH EXPERIENCE

---

### **National Renewable Energy Lab, Golden, Colorado**

*June 2022 - Current*

- *Graduate (Year-Round) Researcher Intern Artificial Intelligence and Energy Efficiency*
- Mentor: Charles Tripp
  - Applying machine learning techniques to renewable energy and energy efficiency for modular network validation and performance
    - \* Research and analyze current trends in modular AI
    - \* Programtically develop software that helps identidy subnetworks of neural networks for optimal validation

### **Cornell University, Ithica, New York**

*June 2022 - Current*

- *Intern: Institute of Biotechnology*
- Mentor: Dr. William Kai Ming Lai
  - Applying Machine Learning approaches to epigenomic datasets for research data generated by the Epigenomics Core Facility
    - \* Developing and integrating new Machine Learning data analysis methods and quality control metrics into the genomics pipeline of the bioinformatics team
    - \* Analyzed and interpreted experimental results, develop new computational methods, and co-author research reports for presentation or publication

### **National Renewable Energy Lab, Golden, Colorado**

*May 2022 - August 2021*

- *Summer Intern: Renewable Resources and Enabling Science Center*
- Mentor: Dr. Nolan Wilson
  - Inverse Bioproduct Design Through Machine Learning and Molecular Simulation
    - \* Using Graph Neural Networks compress molecular polymer structures into a latent feature space that minimizes distances between similar molecules
    - \* We aim to increase ML accuracy and screening using multiple properties to improve the ability to down select new molecules

### **National Renewable Energy Lab, Golden, Colorado**

*June 2021 - August 2021*

- *Summer Intern: Renewable Resources and Enabling Science Center*
- Mentor: Dr. Michael Crowley & Dr. Lintao Bu
  - Optimizing and porting Fortran90 code to NRELs HPC Eagle
    - \* Worked on optimizing polymer physics simulation code using openMP, MPI, and algorithm redesign.
    - \* Observed around a 30 % computational speed up on the code

### **Los Alamos National Laboratory, Los Alamos, New Mexico**

*September 2019 - August 2020*

- *Post Baccalaureate Researcher: Fluid Dynamics and Solid Mechanics (Theoretical Div 3)*
- Mentor: Dr. Luke Van Roekel
  - Working on a team to improve the coverage of testing for the Model for Prediction Across Scales (MPAS) Ocean model
    - \* Developing a workflow and interface to implement test cases
    - \* Creating regression tests for the current MPAS-Ocean model

### **Los Alamos National Laboratory, Los Alamos, New Mexico**

*June 2019 - August 2019*

- *Organizer: Information Science and Technology Institute (ISTI) Rapid Response*
  - Organizer – “Developing Generative Adversarial Networks (GANs) Competition”
  - Developed Python tutorials for writing an MNIST GAN

**Los Alamos National Laboratory, Los Alamos, New Mexico**

*July 2018 - August 2019*

- *Post Baccalaureate Researcher: X-Division Computational Physics, Verification and Validation (XCP-8)*
- Mentor: Dr. Kyle Hickmann
  - Contributed to a Python software infrastructure that simulates multi-physics high-velocity impact models.
  - Studied the strength properties of Beryllium in a Preston-Tonks-Wallace (PTW) model using flyerplate data.
  - Presented work on calibration efforts to characterize Beryllium strength using experimental data from flyerplate impact experiments.
    - \* Implemented a Gravitational Search Algorithm (GSA) to heuristically traverse the parameter space of a subset of default parameters in the PTW model under the BSD-3 License for reproducibility

**Los Alamos National Laboratory, Los Alamos, New Mexico**

*August 2018 - March 2019*

- *Post Baccalaureate Researcher: Intelligent Space Research Division, Space Data Science and Systems (ISR-3)*
  - Developed an XML validation tool to ensure users input a compliant XML schema format for extracting binary satellite data.
  - Collaborated on a team to verify the data returned from legacy satellite system.
    - \* Analyzed and validated tool using Python unit testing
    - \* Cross validated HDF5 file sets to meet IEEE standard of quality

**Los Alamos National Laboratory, Los Alamos, New Mexico**

*June 2018 - August 2018*

- *Summer Intern: Parallel Computing Summer Research Internship*
- Mentors: Dr. Carrie Manore, Dr. Geoffery Fairchild
  - Leveraged data from multiple sources (Satellites, Google health trends, and government data); worked on parallel processing and analysis to predict the spread of Dengue Fever in Brazil
  - Presented work on implementing parallelism to speed up a big data workflow in order to efficiently build statistical models that predict the spread of Dengue fever in Brazil
    - \* Developed an HPC multi-node multi-processor API to distribute software across nodes and processes
    - \* Integrated a large set of diverse input streams into different models via data parallelism for cross model validation.

**Los Alamos National Laboratory, Los Alamos, New Mexico**

*June 2017 - August 2017*

- *Summer Intern: Intelligent Space Research Division, Space Data Science and Systems (ISR-3)*
- Mentor: Dr. Diane Oyen
  - Developed a set of tutorials and walk through on how to integrate user-defined analytic algorithms into the Morphological Analysis of MAterials (MAMA) software.
  - Presented quantifiable analysis of molecular image segmentation
    - \* Worked on an extension of the MAMA program that allows individuals to create their own plug-in and incorporate it into the main software.
    - \* Implemented an algorithm to quantify image data to parse and identify material fragments in a microscopic image

**CUNY Lehman College, Bronx, New York**

*September 2016 - May 2017*

- *Researcher: Lehman College Chemistry Department*
- Mentor: Dr. Tom Kurtzman
  - Developed software to scrape and analyze data from the RCSB Protein Data Bank and identify topographical motifs of bridging water molecules
    - \* Created fast matching algorithms that incorporate bridging water contributions into protein-ligand docking scores which could improve computational pose and binding affinity predictions exponentially.

## - INVITED TALKS

---

### **Career Choices in Computer Science: How to make the right mistakes?**

*Spring 2021*

- *Google Development Group (GDG) - Bronx Chapter*
- Talked about my experience in Academia and Research to undergraduate students interested in pursuing graduate school and industry.

### **Neural Networks and how to make them Fight**

*Spring 2021*

- *Google Development Group (GDG) - Bronx Chapter*
- Presented an interactive tutorial on the basic foundation of Neural Networks, Activation Functions, and GANs

### **An Introduction to Machine Learning**

*Fall 2017*

- *Google Development Group (GDG) - Bronx Chapter*
- Presented my tutorials on how to start writing Machine Learning algorithms in Python using Keras and Tensorflow.
- Presented a tutorial on audio classification using Keras in a live stream example using input from volunteers.

### **Parallel Computing: Why, How and When to Use It**

*Fall 2018*

- *Google Development Group (GDG) - Bronx Chapter*
- Introduced the basics of Slurm management, programming using data and task parallelism, using Python's multiprocessing API to the GDG student monthly meeting.

### **International RSE Day: US-RSE**

*Fall 2021*

- *US Research Software Engineers (USRSE)*

## TEACHING

---

### **CS:2230 - Data Structures and Algorithms**

*Fall 2020 - Fall 2021*

- Teaching Assistant under the University of Iowa Computer Science Department
- Worked collaboratively with students and the professor in order to update the lecture. This was to make the class interactive and work around different forms of learning for students.
- After changing the lesson plans we saw the students retain the class information better and perform better on assignments.

### **Pathway to Student STEM Success Program (PTS<sup>3</sup>)**

*Spring 2017 - Fall 2018*

- Teaching Assistant under the PTS<sup>3</sup> program for a Calculus one course
- Tutored and Mentored Student who were enrolled in the PTS<sup>3</sup> program; this mentorship has continued to today where some of my mentees have moved on to work at companies like Northrop Grumman or graduate schools like the University of Buffalo

### **Introduction to Leadership**

*Fall 2015 - Fall 2018*

- Teacher under the Herbert H. Lehman Leadership Program
- Facilitated a shared safe learning environment for 60 students

- Collaborated with other facilitators in researching, designing, and developing lesson plans that are aligned to the five practices of exemplary leadership

## PRESENTATIONS

---

- J. Gafur**, Workshop, “INTERSECT Research Software Engineers (RSE)”, Princeton University, Workshop, Fall 2022
- J. Gafur**, Lintao Bu, Michael Crowley “Non-equilibrium Molecular Simulations of Polymers under Flow: Saving Energy through Process Optimization”, Energy Efficiency and Renewable Energy (EERE) Advanced Manufacturing Office (AMO) Conference, Summer 2021
- J. Gafur**, Lintao Bu, Michael Crowley “Non-equilibrium Molecular Simulations of Polymers under Flow: Saving Energy through Process Optimization”, Association for Computing Machinery (ACM) Practice & Experience in Advanced Research Computing (PEARC) Conference series, Summer 2021
- J. Gafur**, Jamil Gafur, David Howard Neill-Asanza, Carrie Manore, Geoffrey Fairchild, “Pyaesar: A Multi-Node Multi-Processor API”, Scipy 2020, SciPy Conference 2020, LA-UR-20-21620.
- J. Gafur**, “Julia Embedding for In-Situ Analysis of Atmospheric Modeling”, LANL Student Talks, 03-18-2020.
- J. Gafur**, E. Tourange, K. Hickmann, M. Prime, “Calibration of Flyer Plate Impact Experiments using Particle Swarm Optimization (PSO) Strategies”, American Society of Mechanical Engineers (ASME), 01-18-2019, LA-UR-19-20900.
- J. Gafur**, K. Kempfert, “Forecasting Dengue in Brazil with Time Series Modeling in Parallel”, LANL Student Symposium, 08-03-2018, LA-UR-18-27438.
- J. Gafur**, K. Kempfert, “Nuclear Material Analysis: Developing a Plugin for the MAMA software”, LANL Student Symposium, 07-26-2017, LA-UR-17-26353.

## SOFTWARE PACKAGES

---

- J. Gafur**, E. Tourange, K. Hickmann, Particle Swarm Optimizer: Developer Tools BDS-3 Licensing, Los Alamos National Lab, September 2019
- J. Gafur**, G. Fairchild, C. Manore, Pyaesar BDS-3 Licensing, Los Alamos National Lab, February 2019

## PUBLICATIONS

---

- J. Gafur**, E. Tourange, K. Hickmann, M. Prime, Extending Beryllium Strength Calibration To High Strain-rate Regimes, LA-UR-19-29745.

## COMMUNITY/ACADEMIC REVIEW

---

- **Scientific Computing (SC) 2022 Conference - Artifact Reviewer**
- **Scipy 2022 Conference - Reviewer**
- **Scipy 2020 Conference - Reviewer**
- **STEM-Trek ScienceSlam@SC21 - Judge**