Emily J. Costa

emily.costa.j@gmail.com | +1 (941) 893.7333 | github.com/emilyjcosta5

WORK EXPERIENCE

Oak Ridge National Laboratory **Oak Ridge, TN**
GEM Fellow Intern – Mentor: Dr. Georgia Tourassi *June 2021 – Present*

- Training scalable NLP deep learning models to perform computational phenotyping on clinical text in order to automate cancer report classification for an Exascale Computing Project, CANDLE, in collaboration with the National Cancer Institute
- Profile applications as well as optimize data motion in order to enable highly parallel code deployment and improve performance on state-of-the-art complex HPC node architectures specifically within the Summit and Frontier supercomputers

Northeastern University **Boston, MA**
Graduate Research Assistant – Mentor: Dr. Devesh Tiwari *August 2020 – Present*

- Apply machine learning and statistical inference to analyze I/O patterns in HPC workloads in order to make future trend predictions and improve job scheduling based on probing data of large-scale clustered file systems of various configurations
- Design solutions for identifying and resolving I/O issues affecting production HPC systems and data-intensive jobs through open-source packages and publishing research papers in order to enable system admins to devise new management policies
- First-authored an HPC research paper that is accepted and expected to be published in Supercomputing on November, 2021

Florida International University **Miami, FL**
Undergraduate Research Assistant – Mentor: Dr. Fahad Saeed *September 2019 – February 2020*

- Optimized and revamped file compression for improving I/O performance and reliability in scalable mass spectrometry proteomic applications in order to implement novel methodology thus improving computational and experimental workflow

Oak Ridge National Laboratory **Oak Ridge, TN**
Undergraduate Research Summer Intern – Mentor: Dr. Suhas Somnath *June 2019 – August 2019*

- Developed a computational framework for scaling a Bayesian inference algorithm which integrated in open-source scientific Python packages, pyUSID and pycroscopy, in order to facilitate large-scale spectroscopic image processing and data storage
- Built automated software tests that support continuous integration in order to reinforce package reliability and code quality

PROJECTS

Command my Stocks **Boston, MA**
github.com/emilyjcosta5/cmdmystocks *January 2020*

- Created an open-source Linux command line user interface to easily customize trading bots on Google Cloud Platform then launched an algorithm to trade based on the magnitude of stock price changes in order to make fintech more accessible

Smoky Mountain Conference Data Challenge **Oak Ridge, TN**
github.com/emilyjcosta5/datachallenge2 *July 2019*

- Implemented a convolutional neural network, ResNet-50, to classify electron diffraction patterns on Summit, the world's fastest supercomputer, in order to reduce the workload of traditional compute- and data- intensive spectroscopy applications
- Balanced over- and under-represented crystallographic space groups in over 600 gigabytes of data using dynamic data preprocessing techniques and using SMOTE to generate pseudo-images in order to improve model training and performance

OmiCloud **Boston, MA**
github.com/emilyjcosta5/OmiCloud *March 2019*

- Developed and implemented sophisticated algorithms that rapidly detect surfaces and moving objects using three-dimensional points collected by a Microsoft Kinect in order to identify and monitor a fallen person and alert emergency medical services

EDUCATION

Northeastern University **Boston, MA**
Master of Science Computer Engineering, Cumulative GPA: 3.9/4.0 *August 2020 – Present*

- **Funding & Distinctions:** GEM Full Fellow, Research Assistantship
- **Relevant Coursework:** Computer Architecture, Computer Systems, Data Structures, Machine Learning

Florida International University **Miami, FL**
Bachelor of Science Mathematics and Computer Science, Cumulative GPA: 3.5/4.0 *August 2017 – July 2020*

ADDITIONAL SKILLS & INTERESTS

Technical: BASH, C/C++, Deepspeed, Docker, Git, HuggingFace, Java, Julia, Linux, Pandas, Python, Pytorch, R, Scikit-Learn
Interests: Cooking, Levantine Arabic, Mathematical Reasoning, Philosophy, Piano, Sailing, Theology