Sonia Sharapova

OBJECTIVE

Computer Science MS student specializing in High Performance Computing seeking to advance AI/ML model development and optimization through expertise in parallel computing and deep learning model development.

CONTACT

(408) 797-8495

sonia.sharapova@yahoo.com

linkedin.com/in/sonia-sharapova

github.com/sonia-sharapova

EDUCATION

Sept. 2023 - March 2025
The University of Chicago
M.Sc. Computer Science
Spec. High Performance Computing
– Chicago, IL

Relevant Coursework:
HPC, Parallel Programming, Cloud Computing, Algorithms, Unix Systems, Python

Sept. 2019 - July 2023
McGill University
B.Sc. Computer Science and Biology
– Montreal, QC, Canada

Relevant Coursework:
Machine Learning, Computer Vision, Linear Algebra, Statistics

Awards:
Meakins-Christie Laboratory Studentship Competition
- *Maximum stipend for research.*

SKILLS

Programming Languages:
Python, C, Java, Go, R, Bash

Web Development:
HTML, CSS, JavaScript

Scientific Computing / Data Science:
MATLAB, NumPy, Pandas, Scanpy, OpenCV, Jupyter Notebook, Data Visualization

Machine Learning and Parallel Programming:
PyTorch, Keras, TensorFlow, CUDA, Open MPI, OpenMP, Hugging Face

Cloud and DevOps:
AWS, Git, Linux/Unix

PROFESSIONAL EXPERIENCE

Data Analyst Intern
May - August 2024
Elevance Health - Chicago, IL

- Led the transition from PyTorch to the Hugging Face ecosystem for an existing custom GPT-2-based generative model (MediClaimGPT), enabling standardized deployments and improved model versioning.
- Created tutorials and delivered technical presentations to educate my team on the Hugging Face environment.
- Maintained model performance through the migration, preserving its ability to generate clinically plausible synthetic data for classification, population analytics, and predictive modeling tasks.

Student Researcher
May 2021 – May 2024
McGill University Health Center - Montreal, QC, Canada

- Implemented and trained a novel hybrid machine-learning model which combines an autoencoder with a multilayer perceptron. This model resulted in higher accuracies predicting patient cancer types compared to standard models.
- Computed a customized autoencoder loss function that resulted in gene marker identification.
- Applied sensitivity analysis and hyperparameter tuning to iteratively improve training model performance.
- Analyzed data and performance results with MATLAB and Python's Matplotlib; visualized gene expression with the Python toolkit Scanpy to find gene clusters associated with corresponding cell types.
- Analyzed the correlation between muscle features and respiratory failure in Cystic Fibrosis patients using data fitting, clustering methods, and comparative model analysis.

Cardiac Deformation Analysis Project
December 2024
The University of Chicago - Chicago, IL

- Developed parallel processing systems in both Go and Python for analyzing cardiac motion in medical imaging sequence, comparing the efficiency of both environments.
- Implemented pipeline and work-stealing parallel architectures in Go utilizing native concurrency primitives (goroutines and channels), achieving a 3.22x speedup with 6 workers.
- Created equivalent Python system using scientific computing libraries (OpenCV, NumPy, pydicom), optimizing the codebase for efficient DICOM file processing and feature detection.
- Analyzed performance trade-offs between languages: Go's superior concurrent performance versus Python's rich ecosystem of scientific libraries.