Md Shihab Shahriar Khan

Stack Overflow: /u/4553309/
GitHub: /shihab-shahriar
Email: khanmd@msu.edu
Phone: (+1) 517-775-9450

Experience

Michigan State University — MI, USA
Research Assistant II — Mar. 2022 – Present

- Developing a performance-portable simulation framework for particulate flows, designed for seamless scaling across multi-node systems and on-node with GPU acceleration.
- Tech stack includes C++, CUDA, OpenMP/MPI, Kokkos, Python

Samsung R&D Institute Bangladesh — Dhaka, Bangladesh
Machine Learning Engineering Intern, Advanced Research Group — Jan. 2018 – Jun. 2018

- Worked primarily in computer vision, particularly image classification, inference on edge etc.
- Technologies included tensorflow/keras, numpy, pandas etc.

Singularity Corporation — Dhaka, Bangladesh
Software Engineer Intern — Jan. 2016 – Mar. 2016

- Worked on a time tracking app- to track work hours, make payrolls, and bill clients
- Part of the backend development team, used technologies like Django, MySql etc.

Publications

1. **Shihab Shahriar Khan**, Nishat Tasnim Niloy, Md. Aquib Azmain and Ahmedul Kabir. "Impact of Label Noise and Efficacy of Noise Filters in Software Defect Prediction". *International Conference on Software Engineering and Knowledge Engineering (SEKE)* KSIR Virtual Conference Center, USA, 2020.

2. **Shihab Shahriar Khana**, Bryce Palmer , Christopher Edelmaier , Hasan Metin Aktulga. "OpenRAND: A Performance Portable, Reproducible Random Number Generation Library for Parallel Computations", **Under Review on SoftwareX**

Education

- **Michigan State University** — MI, USA
  PhD in Computer Science & Engineering — Aug. 2021 – Present
  - CGPA: 3.92/4.00 (completed 18 credits)
- **Institute of Information Technology, University of Dhaka** — Dhaka, Bangladesh
  M.S. in Software Engineering — Jan. 2019 – Dec. 2020
  - Thesis Title: "Noise Robust Classification Using Instance Level Analysis"
  - CGPA: 3.83/4.00
- **Institute of Information Technology, University of Dhaka** — Dhaka, Bangladesh
  B.S. in Software Engineering — Jan. 2015 – Dec. 2018
  - CGPA: 3.68/4.00

Technical Skills

Machine Learning and Data Science
Python, Pytorch, Numpy/Pandas, R, Scipy, Scikit-learn, SQL, Tensorflow

High Performance Computing
C/C++, CUDA, Kokkos, OpenMP/MPI, Trilinos, CMake/Make, Paraview/VTK

Common Tools and Others
Data structure and Algorithms, Git, Linux, CI/CD, WebGPU, Sphinx/Doxygen, Haskell, Java

Projects

OpenRAND (Link)
Reproducible random number generation for parallel computations (Lead Developer) Sept. 2023 – Present

- An open-source C++17 library aimed at facilitating reproducible scientific research through the generation of reproducible, parallel random number streams
- Cross-Platform, designed to work seamlessly across various software and hardware platforms, including GPUs.
- Performance, It is as fast and often faster than native libraries like libstdc++, Nvidia's Curand or rocRAND.
- Statistical Robustness, design choices and built-in tests ensure that the random number streams generated by OpenRAND are statistically robust, with no discernible patterns.

Scikit-clean (Link)
A python ML library for classification in the presence of label noise. Dec. 2019 – Dec. 2021

- A collection of algorithms for detecting and handling label noise
- Tools to simulate artificial noise, create complex pipelines and evaluate them
- scikit-learn API compatible- all scikit-learn's building blocks can be seamlessly integrated into workflow
- Equipped with units tests, extensive documentation, CI pipeline and available in pypi.

DeepPaint (Link)
Deep learning based computer vision tool to automatically colorize and stylize paintings Jul. 2018 – Nov. 2018

- Colorizes a sketch using color hints.
- Stylizes a colored painting along the style of any given painting.
- Implemented using python, pytorch, pyqt.

sTorrent
A Bittorrent client implemented completely from scratch Jan. 2017 – Jun. 2017

- Fully functional, implements all key parts of Bittorrent protocol
- Implements DHT (distributed hash table) to support magnet links.
- Implemented only using standard python library.

Activities

Open Source
Several code-level pull requests accepted in projects like Kokkos, Spack, Imbalance-learn, and MLAlgorithms

Technical Writing Samples

- **StackOverflow**: Currently with 5k "reputation", mostly in Python, Scikit-learn, Pytorch and Algorithm tags.
- **Article**: "CUDA vs ROCm: A Case Study"
- **Article** in Towards Data Science titled "An Introduction to Classification Using Mislabeled Data".