

JOYMALLYA CHAKRABORTY

<https://joymallya.github.io/> ◊ Github: github.com/joymallya
2359 Champion Ct, NC 27606

Email : jchakra@ncsu.edu
Mobile : +1-919-633-2503

EDUCATION

Ph.D. in Computer Science

North Carolina State University, Advisor - Dr. Tim Menzies (timmm@ieee.org)

Aug. 2017 – March 2022
Raleigh, NC

BE in Computer Science

Jadavpur University

Jun 2011 – Jun 2015
Kolkata, India

TECHNICAL SKILLS

General Expertise: Software Engineering, Machine Learning, Deep Learning, Data Science, Compiler, Program Analysis, Optimization;
Programming: Proficient: C/C++/Java (7+ years), Python(5+ years). Working Knowledge: R, AngularJS, CUDA, NodeJS;
DevOps: Docker, Kubernetes, AWS, Ansible, Vagrant, Travis, Jenkins.

WORK EXPERIENCE

Amazon

Applied Scientist Intern, Abuse Detection

Seattle, WA

May 2021 - Aug. 2021

- Worked as a member of the Abuse Prevention Science team to find emerging abuse on Amazon Product Detail Page.

Languages: Python, SQL; **Frameworks:** AWS, Cradle, Deep Learning

IBM Research

Ph.D. Research Intern, System Design

Yorktown Heights, NY

May 2020 - Aug. 2020

- Worked on State Management & Persistence in Mono2Micro which converts Monolith applications to Microservices.

Languages: Java; **Frameworks:** Kubernetes, Redis, RabbitMQ

Intel Corporation

Software Engineering Research Intern, Deep Learning

Bellevue, WA

May 2019 - Aug. 2019

- Worked on computational graph optimization and post-training quantization of CNN models.

Languages: C++, Python; **Frameworks:** Onnxruntime, Ngraph, TensorFlow, **Model:** Resnet50

Intel Corporation

Software Engineering Research Intern, Compiler Optimization

Bellevue, WA

May 2018 - Aug 2018

- Explored optimization opportunities of .NET Core Garbage Collection and implemented PoC (Proof of Concept) prototypes.

Languages: C++, .NET; **Frameworks:** .NET Framework, VTune

TCG Digital

Software Developer, Java & AngularJS

Kolkata, India

June 2015 - July 2017

- Worked in a Business Intelligence Software to retrieve, analyze, transform and report business related data.

Languages: Java, Bootstrap, JQuery, Angular JS, Node JS, Require JS, C3 JS, D3 JS; **Frameworks:** Elasticsearch

RESEARCH PROJECTS

Reducing Training Time of DNN Models

LexisNexis funded project in the RAISE Lab

Raleigh, NC

Jan 2021 - Present

- Using Local interpretable model-agnostic explanations (LIME) to reduce training time for DNN models using NetFlow data.

Fairness Aware Classification

NSF funded project in the RAISE Lab

Raleigh, NC

Feb 2019 - Present

- Generated fairer results for various learners using fairness as a goal in pre-processing & in-processing.
- Used ML model explanation to find underlying causes of unfairness.

Measuring Effects of Heroism Using Socio-technical Network

NSF funded project in the RAISE Lab

Raleigh, NC

Aug 2018 - Dec 2018

- Studied the effect of heroism on 1100+ GitHub projects.
- Found out that hero developers are far less likely to introduce bugs into the codebase than their non-hero counterparts.

Gender Differences and Bias in Open Source: Pull Request Acceptance of Women versus Men

NSF funded project in the DLF Lab

Raleigh, NC

Dec 2017 - May 2018

- Analyzed GitHub Pull Requests to find out the root causes of gender bias against women.

Enabling Flexible Task Assignment On GPU Through SM Centric Program Transformations

Class project.

Raleigh, NC

Aug 2017 - Dec 2017

- Developed a source to source translator that can perform SM centric program transformation with help of Clang's LibTooling, ASTMatcher, LoopRewriter.

All Possible Spanning Tree Generation of a Simple, Symmetric, Undirected Graph

Research project in Calcutta University

Kolkata, India

Aug 2016 - July 2017

- Implemented C code version of twelve algorithms proposed by several researchers and compared those based on CPU time.
- A complete new algorithm of spanning tree generation has been constructed.

Diagnostic Evaluation using Linguistic Checkpoints For Machine Translation (DELIC4MT)

Research project in Jadavpur University

Kolkata, India

Aug 2014 - June 2015

- Used PoS taggers and word aligner to represent word alignment, KYOTO Annotation Format (KAF) to represent textual analysis. Later, Kybots were used to extract linguistic phenomena.

Secure Authorization System based on Finger Vein Identification

Research project in Jadavpur University

Kolkata, India

Feb 2012 - Mar 2013

- Extracted ROI (Region of Interest) first from images, then used Intensity Normalization techniques. Later, Contrast Limited Histogram Equalization was done followed by Average Filtering.

SELECTED PUBLICATIONS

- [1] Chakraborty J., Majumder S., Menzies T., "Bias in Machine Learning Software: Why? How? What to do?". In **Foundations of Software Engineering (ESEC/FSE), 2021**. Online: <https://arxiv.org/abs/2105.12195>; **ACM SIGSOFT Distinguished Paper Award Winner**
- [2] Chakraborty J., Majumder S., Yu Z., Menzies T., "Fairway: A Way to Build Fair ML Software". In **Foundations of Software Engineering (ESEC/FSE), 2020**. Online: <https://dl.acm.org/doi/10.1145/3368089.3409697>;
- [3] Chakraborty J., Peng K., Menzies T., "Making Fair ML Software using Trustworthy Explanation". In **International Conference on Automated Software Engineering (ASE), 2020**. Online: <https://dl.acm.org/doi/10.1145/3324884.3418932>;
- [4] Chen J., Chakraborty J., & Menzies T., "Predicting Breakdowns in Cloud Services (with SPIKE)". In **Foundations of Software Engineering (ESEC/FSE), 2019**. Online: <https://dl.acm.org/doi/abs/10.1145/3338906.3340450>;
- [5] Imtiaz N., Middleton J., Chakraborty J., "Investigating the Effects of Gender Bias on GitHub" In **International Conference on Software Engineering (ICSE), 2019**. Online: <https://ai.google/research/pubs/pub47860>;
- [6] Chakraborty M., Chowdhury S., Chakraborty J., Mehera R., Pal R., "Algorithms for generating all possible spanning trees of a simple undirected connected graph: an extensive review" In **Complex & Intelligent Systems (Springer), 2018**. Online: <https://link.springer.com/article/10.1007/s40747-018-0079-7>;