1735 Woodland Ave, Apt 62
East Palo Alto, CA 94303

Saubhik Mukherjee

Mobile: (470) 313-0534
saubhik.mukherjee@gmail.com

Employment

Senior Software Engineer — **SambaNova Systems, Palo Alto** — **May 23, 2022 - Present**
Machine Learning (ML) Compiler

- Build MLIR and LLVM-based compiler layers and compiler tools to transform, optimize, debug, and execute ML models on proprietary ML accelerator architectures.
- Build scalable and high-quality production compiler infrastructure using well-established and emerging techniques and push the boundaries of compiler design.
- Develop, maintain, and debug compiler optimization algorithms on ML graphs and add compiler support for new hardware architectures.
- Analyze and improve compile-time and run-time performance across multiple AI hardware architectures and ML frameworks, such as TensorFlow and PyTorch, to support new state-of-the-art training and inference.
- Collaborate with ML researchers and engineers to guide compiler development for future ML trends.
- *Tools*: C++, (Clang), MLIR, LLDB, cmake, ninja, gperf, tmux, neovim, ctags, clangd, Synopsys VCS

Graduate Research Assistant — **Georgia Tech, Atlanta** — **Aug '21 - May '22**

- Ported Facebook's QUIC implementation, mvfst, to rely on the efficient kernel-bypass network stack (threading & socket) APIs provided by MIT's Shenango and achieve low tail latency & increase CPU efficiency.
- *QuicNIC*: Offloaded GSO & crypto (encryption, decryption) to a dedicated CPU core to obtain record QUIC throughputs (x5). *Skills*: QUIC, caladan, mvfst, folly, fizz, profiling prod C++ codebase, CPU FlameGraphs.

Linux Contributor, Intern — **Google Summer of Code** — **May '21 - Aug '21**

- Analyze and fix race condition bugs in the Linux kernel 5.4 device drivers based on software verification static analysis tool, Klever. Part of *Linux Standards Base* & *Linux Driver Verification*.
- Accepted patches to kernel mainline. *Skills*: Linux kernel development, C.

Lead Software Engineer — **Alpaca, Tokyo, Japan** — **(2 years) Jan '19 - Jan '21**

- Agile team lead of 3 engineers & Release Manager developing prod systems for delivering JP & US stock predictions to hedge funds, asset management firms through back-end batch systems & web-apps.
- Lead dev & design of full-stack web-service (JS, Python, Flask, PostgreSQL) deployed to prod from scratch withing 3 months; led 10% annual revenue growth; involved in code reviews, schema design & prod deploys.
- Dev & design of prod ML pipeline & ML model dev for stock price predictions using CNN, NN arch search (PyTorch, Pandas, NumPy)

Data Scientist — **Anheuser-Busch InBev, Bangalore** — **(1.5 years) Jun '17 - Dec '18**

- Developed pricing conjoint modelling pipelines for revenue mgmt across business units with UI dashboards.
- Awarded *Innovation Impact Award* among 200 new hires for developing brand/SKU optimization techniques, aggregating regional conjoint studies; lead 4 data scientists. *Tools*: R, RShiny, Python, Jupyter.

Education

Atlanta, GA — **Georgia Institute of Technology** — **Jan '20 - May '22**

- **Master of Science in Computer Science** with *Systems Specialization*, May 2022. **GPA: 4.0**
- *Courses*: Operating Systems; Computer Architecture; Compilers; Networks; Distributed Systems; Databases; HPC; Algorithms.

Kolkata, IN — **Indian Statistical Institute** — **Jul '15 - Jun '17**

- **Master of Science in Quantitative Economics**; full scholarship & monthly stipends
- *Relevant Courses*: Optimization; Game Theory; ACM-ICPC regional qualifier.

Chennai, IN — **Chennai Mathematical Institute** — **Aug '12 - Apr '15**

- **Bachelor of Science in Mathematics and Computer Science**, Innovation in Science Pursuit Scholar
- *Courses*: Algorithms; Programming Languages; Discrete Math; Theory of Computation; Logic.

Projects

- *Xen's Credit Scheduler* Credit scheduler C implementation in a user-level threads library.
- *TinyFile* Shared memory-based file compression service with client library in C.
- *Distributed KVS* Distributed key-value store using XML-RPC with partitioning, replication & consistency.

Languages and Technologies

- **About 10k lines**: C++, Python, C, Java; **About 500 lines**: JS, HTML, CSS
- Linux, git, tmux, vim, ctags, LSPs (clangd), docker, k8s, AWS, Auth0, CircleCI, DataDog, ArgoCD