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

# Saubhik Mukherjee

+1 (470) 313-0534
saubhik.mukherjee@gmail.com

## Employment

**Senior Software Engineer** — **SambaNova Systems, Inc.** — **May 23, 2022 - Present**
***Machine Learning (ML) compiler*** — **Palo Alto, CA, USA**

- 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.

**Software Engineer** — **AlpacaJapan, Co. Ltd.** — **Jan 1, 2019 - Jan 8, 2021**
***Machine Learning (ML) systems*** — **Tokyo, Japan**

- Design, develop, maintain, and test live production software systems for delivering stock price predictions.
- Collaborate with data science & engineering team to integrate different software systems and deploy and upgrade ML models in live production financial forecasting software.
- Handle installation & maintenance of new data sources and develop the data platform used for ML model R&D.
- Collect and document client requirements for future releases and make extensible and robust software design decisions for developing server & client web applications; responsible for 10% annual revenue growth.
- Manage software releases with an agile mindset and develop workflows for fast production recovery in case of failures.
- Driving innovation by evaluating new technologies, original financial data sources, and recent research papers that add value to Alpaca's products.
- *Tools*: Python, React, JavaScript, Flask, PostgreSQL, SQLAlchemy, Alembic, Kubernetes, Docker, PyTorch, Pandas, NumPy, SciPy, Luigi, CircleCI, Argo CD, Auth0, Datadog.

**Data Scientist** — **Anheuser-Busch InBev** — **Jun 19, 2017 - Dec 28, 2018**
***ML in pricing research*** — **Bangalore, India**

- Develop machine learning models to estimate ABInBev's market share and revenue in different pricing scenarios of beer SKUs across multiple countries, using both R statistical programming language and Python.
- Conduct extensive experiments to determine the significant variables in ML models and create automated scripts to replicate the process, using Keras, TensorFlow.
- Create pricing conjoint survey questionnaires and handle data management and pre-processing using customized scripts and workflows; used dplyr, tidyr.
- Interact and collaborate with business heads in different countries to include different pricing scenarios in conjoint based on the business requirements and present the pricing analysis results for business actions; used ggplot2.
- Develop various optimization algorithms based on pricing analysis results to maximize the business objective, such as market share or revenue; used nloptr.
- Create UI dashboards that display conjoint analysis results for business to gain actionable insights, using Shiny, RStudio.

## Internships & Research

**Graduate Research Assistant** — **Georgia Institute of Technology** — **Aug 23, 2021 - May 7, 2022**
***Networked Systems*** — **Atlanta, GA, USA**

- Developed *QuicNIC*, a software NIC to accelerate the QUIC stack, to move segmentation, pacing, and encryption from the application to a dedicated core implementing software NIC functionality; achieved TCP-like single-connection throughput.
- Ported Meta's production QUIC (network protocol) implementation, mvfst (& dependencies folly, fizz), to rely on the efficient kernel-bypass network stack (custom threading & socket libraries) APIs provided by MIT's Shenango (caladan) and achieve low tail latency & increase CPU efficiency; involved CPU profiling using flame graphs.
- This research was supervised by professors Ahmed Saeed and Mostafa Ammar.

**Open Source Contributor** — **Google Summer of Code** — **Jun 7, 2021 - Aug 23, 2021**
***The Linux Foundation***

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

## Education

**Master of Science in Computer Science, *Systems Specialization*** — **Georgia Institute of Technology, Atlanta, GA, USA** — **Jan 6, 2020 - May 7, 2022**

- *Major*: Computer Science, *Concentration*: Computing Systems. *Overall GPA*: 4.00. Transcript link.
- *Courses*: High Performance Computer Architecture, Computer Networks, Compilers & Interpreters, Advanced Operating Systems, High Performance Computing, Database Systems Concepts & Design, Distributed Computing, Introduction to Graduate Algorithms, Datacenter Networks & Systems.

**Master of Science in Quantitative Economics** — **Indian Statistical Institute, Kolkata, India** — **Jul 2015 - Jun 2017**

- Passed in First Division with Distinction. Transcript link.
- *Courses*: Microeconomics I, Game Theory I, Statistics, Computer Programming and Applications, Basic Economics, Microeconomic Theory II, Macroeconomic Theory I, Econometric Methods I, Time Series Analysis and Forecasting, Game Theory II, Macroeconomics II, Econometric Applications I, Individual and Collective Choice, Political Economy, Econometric Applications II, Economic Development, Bayesian Econometrics, Auction Theory, Optimization Techniques.

**Bachelor of Science (Honours) in Mathematics & Computer Science** — **Chennai Mathematical Institute, Chennai, India** — **Aug 3, 2012 - Jul 25, 2015**

- *Cumulative GPA*: 06.93 (out of 10). Transcript link.
- *Courses*: Algebra I, Calculus I, Classical Mechanics I, English, Introduction to Programming, Advanced Programming, Algebra II, Calculus II, Discrete Mathematics, Probability Theory, Environment Course, Algebra III, Calculus III, Design & Analysis of Algorithms, Real Analysis, Theory of Computation, Complex Analysis, Differential Equations, Programming Language Concepts, Topology, Differential Geometry, Algebra IV, Introduction to Logic, Commutative Algebra, Optimization Techniques, Economics, Electrodynamics I, Finance, Representation Theory.