Alapan Chaudhuri

alapan.chaudhuri@research.iiit.ac.in banrovegrie.github.io

Education

IIIT Hyderabad **2019 to 2024**
B.Tech. and M.S. in Computer Science and Engineering *Hyderabad, India*

Experience

Instabase **since June 2024**
Software Engineer, Machine Learning *Bangalore, India*

- Developing memory-augmented approaches for continual learning in enterprise search, enhancing contextual awareness and enabling personalized, adaptive retrieval systems.
- Improved search accuracy for smaller corpuses from 87.6% to 95.1% by integrating dynamic KGs, LTC-RNNs, and entity embeddings with long-term memory mechanisms.

Instabase **May 2023 to Aug 2023**
Software Engineer, Intern *Bangalore, India*

- Built a language model interaction and prompting framework to improve document information extraction in AI Hub with 0.98 F1 score for invoices and 0.85 F1 score for 10-K filings.
- Implemented dilated attention to augment Large Language Models (LLMs) with bigger context window sizes and led to a reduction in training costs by 82% for a 6.7B parameter model. This was possible by swapping standard attention layers with dilated attention allowing for sub-quadratic (almost linear) computational complexity.

Centre for Quantum Science and Technology **2022 to 2024**
Research Assistant *Hyderabad, India*

- Worked on the computational complexity of predicting avoided crossings and solving combinatorial optimization problems using adiabatic quantum computation.

Ayers Lab, McMaster University **2022**
Research Assistant *Hamilton, Canada*

- Worked with Dr. Paul Ayers on optimization algorithms for solving the positive semi-definite Procrustes problem and extending the Procrustes python library. Contributed under Open Chemistry in Google Summer of Code 2022.

Skills

Python, C/C++, Haskell, Coq, JavaScript, Racket, MySQL, HTML/CSS, TensorFlow, PyTorch, Cirq, Qiskit, React, Q#

Awards

- **ICPC World Finalist**: Received honourable mention at the 46th World Finals. Placed 2nd (India) at the Asia West Finals 2022 . ACM ICPC is the oldest, largest & most prestigious programming contest in the world.
- **QHack 2022 Winner**: 1st place out of 800+ teams in the Quantum Chemistry Challenge at QHack 2022 by Xanadu.

Publications

- Unstructured Adiabatic Quantum Optimization: Optimality with Limitations (under review PRX Quantum)
- Classifying CELESTE as NP-Complete (CST 2022)

Projects

Fuzzy Logic | *Finance, Statistics, Python* **2024**

- Designed and managed equity-focused trading strategies. Created a library for asset pricing and portfolio optimization, exclusively focused on NSE/BSE, systematically monitoring and evaluating data for 247 tickers.

Lockness | *Python, PyTorch, Machine Learning, Cryptography* **2024**

- Developed a platform for privacy preserving machine learning using secure multi-party computation and oblivious transfer. This allows model developers find, and securely train on datasets that were previously unavailable due to privacy concerns, without leaking neither the data nor the model weights to any party.

SmolGPT | *Python, PyTorch, NLP, Transformers* **2023**

- Developed a character-level auto-regressive transformer model with temperature control and input data filtration for generating multilingual text messages.

Racket Compiler | *Racket, Functional Programming, Compiler Design* **2022**

- Developed a nano pass compiler for a subset of the racket language (supporting vectors and loops as well). Optimized register allocation with greedy graph coloring and recursive calls with tail call optimization.

PauliZee | *Qiskit, Hamiltonian Simulation* **2022**

- Engineered a quantum simulation framework with optimized Trotter methods exceeding the performance of Qiskit's default implementation on 7-qubit IBM systems.

Mariam: a Unix shell | *C/C++, Linux, Operating Systems* **2021**

- Implemented a Unix shell with inter-process communication, custom signal handling, I/O redirection, as well as error handling using raw Linux system calls.

Christine | *Python, NLTK, Google Cloud Platform* **2020**

- Created a discord bot that moderates online harassment along with toxicity and depressive behavior. 1.6 million tweets were used for constructing and calibrating a scale to measure depression from text messages.