

# Inesh Chakrabarti

858-925-3059 | [inesh33@g.ucla.edu](mailto:inesh33@g.ucla.edu) | [linkedin.com/in/inesh-chakrabarti](https://www.linkedin.com/in/inesh-chakrabarti) | [github.com/beesfleas](https://github.com/beesfleas)

## Education

### University of California, Los Angeles

Los Angeles, CA

*B.S. Electrical Engineering, M.S. Electrical Engineering*

*Expected Graduation – June 2027*

- **Graduate GPA:** 4.0 / 4.0, **Undergraduate Major GPA:** 3.8 / 4.0
- **Coursework:** Large Scale Data Scraping Algorithms, Convex Optimization, Deep Learning, Software Engineering, Embedded Systems, Computer Architecture, GPU Microarchitectures, Numerical Computing, Stochastic Systems, DSP, Communications, Speech and Image Processing, Machine Learning, Feedback Control, Signals and Systems, Probability and Statistics
- **Societies:** American Nuclear Society (President), IEEE-Eta Kappa Nu (Mentorship Chair)

## Experience

### UCLA Lin Yang Research Group

*February 2025 - Present*

- *NoWag: A Unified Framework for Shape Preserving Compression of Large Language Models*  
Lawrence Liu, **Inesh Chakrabarti**, Yixiao Li, Mengdi Wang, Tuo Zhao, Lin F. Yang  
Publication accepted to **COLM** and **ICLR SLLM Workshop**
- Built dequantization/inference kernels in C (CUDA) for parallelization over multiple GPUs while using **48x less calibration data** and maintaining performance against SOTA VQ methods
- Implemented Trellis Quantization and benchmarking in Python for NoWag, a set of shape-preserving pruning and quantization algorithms for LLMs

### UCLA Complex Networks Group (Paid Student Researcher)

*February 2022 - June 2024*

- Implemented High Frequency Oscillation Detector using Variational Autoencoder for neural signals
- Processed and visualized neural spike data using Python and MATLAB to demonstrate correlation between individual neural spikes and character recognition from animation

## Projects

### Reinforcement Learning Hearts

*September 2024 - January 2025*

- Created RL agent for Hearts using **Counterfactual Regret Minimization** and **Monte Carlo Tree Search** that reaches approximate Nash Equilibrium and superhuman performance.
- Enhanced the Hearts project with a Tkinter UI and collaborated in a 3-person team, providing a **real-time interface** allowing for physical gameplay simulation via computer vision.

### Algorithmic Trading Bot

*March 2023 - May 2023*

- Developed and deployed ARIMAX-based trading bot utilizing time-series analysis to optimize trading strategies and participate in **live trading** against other bots and consumers.
- Created portfolio optimization bot with Kernel Density Estimation, achieving a **1.4 Sharpe Ratio** and **highest profit** in competition.

### EEG and News Classification

*January 2023 - March 2023*

- Developed a **pipeline** for EEG data analysis with wavelet transform pre-processing and model distillation to predict human movement using transformer, LSTM, and CNN models.
- Experimented with Kmeans, HDBSCAN clustering, NMF, and KNN with Pearson Correlation Recommender Systems
- Constructed an end to end pipeline for news classification using Natural Language Processing (RoBERTa, BOW, TFIDF, SVM, etc.) while leveraging cuML and cuDF for GPU optimization.

### Habit-Tracking IoT Tamagotchi with Web App

*Jan 2022*

- Built backend of web app with MongoDB with login authentication logging users' schedules with an IoT device registering task completion.
- Programmed ESP32 to update device's expressions in real-time based on schedule adherence.

## Skills

- **Programming Languages:** C, C++, Python (NUMBA, PySpark, Matplotlib, PyTorch, Pandas, Keras, Tensorflow), Triton, SQL, x64, C#, Java, MATLAB, R, JavaScript
- **Tools:** Docker, Git, LangGraph, MongoDB, LTSpice, GDB, Unix Shell, CUDA, OpenMP, Joblib, Django, NVIDIA Nsight Compute, Apache Spark, Fuzzing (AFL), CI/CD, Agile Development