Hoang Chu

1050 North Mills Avenue, Claremont, CA | ceric9817@gmail.com | (909)-407-6637 | linkedin.com/in/hoangchu2001

EDUCATION

Pitzer College Claremont, CA
Bachelor's Degree, **Joint Computer Science and Mathematics (Honors)** Expected May 2024

- **GPA**: 3.7 / 4.0 (Dean's List)
- **Teaching Assistant (for 1 year)**: Data Structures (A), Intermediate Probability (A), Intermediate Linear Algebra (A-)
- **Coursework:** Algorithms, Machine Learning, Statistical Linear Models, Combinatorics, Computability and Logic.

AWARDS

- **Math**: 1st Prize: Vietnam Mathematical Olympiad, Silver Medal: Asian Pacific Mathematical Olympiad, 3x AIME.
- **Programming** (college started): Winner - SIG Coding Challenge, Meta Hacker Cup and Google Code Jam Round 2.

RESEARCH EXPERIENCE

University of Southern California - Research Scholar in Operations Research & Data Science June 2023 - Current
"Nonlinear Optimization of Return on Investment for Facility Costs under Probabilistic Utility Models" *Statistics*

- Devised, proved, and tested my proposition: changing from adding to multiplying a deterministic and a random error term transformed the cost optimization function from non-convex to solvable under any uniformly generated datasets.
- Proposed, customized, and implemented a Multi-Cut Approximation algorithm that well-approximated the solution of this NP-Hard problem while performing 5 to 10 times faster than a solver's traditional local search heuristic algorithm.

Harvey Mudd College - Independent Research January 2022 - May 2022
"Explore Constraints on Unitary Recurrent Neural Networks" (**paper**) *Mathematics of Topic Modeling*

- Proved that weight matrices within the unitary group can prevent uRNN's vanishing and exploding gradients problem.
- Developed a sample Unitary Recurrent Neural Network in Python with self-designed update rule based on the previous proof which demonstrated the model's practicality in Language Modeling tasks.

WORK EXPERIENCE

Meta May 2022 - August 2022
Engineering Intern (received return offer) *Python, SQL, Numpy, Scikit-learn, PyTorch, Hadoop (MapReduce)*

- Proposed and delivered a topology-preserving Dimension Reduction algorithm achieving a 93% parameter reduction and improving 2% in predicting accuracy for **100,000,000+ market images** using Meta AI's CommerceMM model.
- Designed a model-agnostic Window Selection algorithm partially addressed Vision Transformer's exponential scaling issue and facilitated a comprehensive performance analysis and comparison of the model with Meta AI's ResNet50.
- Utilized Directed Acyclic Graph data structure to accelerate the automation of generational data workflows in batches.

CoHost.ai (Seed-stage Startup) June 2021 - August 2021
Engineering Intern (only intern in the company) *C++, Linux Development, MongoDB, gRPC, Kubernetes*

- Deployed a fault-tolerant Message Queue with Inter-Thread Communication method which prevents message loss if the program crashes, supports any data types as inputs, streamlines ownership, and leaves no copies of messages sent.
- Architected the system design for the company's new Chatbot and designed test suites reaching **100% code coverage**.

PROJECTS

LendingClub Economic Risk Assessment | Citadel Datathon Winner *R, Regression Analysis, Big Data*

- Performed statistical analysis across **multiple datasets totaling 26,000,000 rows** and discovered that LendingClub had failed to detect new borrowers who'd defaulted and altered their personal information to manipulate interest rates.
- Developed a k-NN forecasting model with 94% accuracy for classifying and predicting new borrowers' interest rates.

Anomaly Detection for Tradeweb's Financial Market Data | p-AI club's sponsored project *Machine Learning*

- Delivered production-graded regression and SVM algos predicting latency spikes in bond prices with 91% accuracy.

Python to Java Compiler in Android Environment | **MLH Fellowship** Open-Source Contributor *Java OOP, Systems*

- Wrote helper functions that remove redundant bytecodes and avoid crashes when translated variables jump addresses.
- Refactored **1200+ lines of code** in the team's analytic tool which monitors the compiler's latency and memory usage.