# Hoang Viet Chu

1050 North Mills Avenue, Claremont, CA | hchu88347@gmail.com | (909)-407-6637 | linkedin.com/in/hoangchu2001

## EDUCATION

**Pitzer College**, a member of the Claremont Consortium — Claremont, CA
**Bachelor's Degree, Joint Computer Science and Mathematics (Honors)** — Expected May 2024

- **GPA**: 3.7 / 4.0 (Major GPA: 3.9 -- Cross-registered Math and CS classes at Harvey Mudd College)
- **Teaching Assistant (for 1 year)**: Data Structures and Algorithms (C++), Intermediate Probability, Abstract Algebra
- **Classes**: Machine Learning, Linear Models, Logic, Computer Systems (OCW), Computer Graphics, Combinatorics

## AWARDS

- **Math**: 1st Prize: Vietnam Mathematical Olympiad, Silver Medal: Asian Pacific Mathematical Olympiad, 3x AIME
- **Coding** (0 high school exp.): 1st Place: Citadel Datathon, Round 3: Google Code Jam (C++), Winner: SIG Challenge

## RESEARCH EXPERIENCE

**University of Southern California** — June 2023 - August 2023
Research Intern - Scholar in Operations Research and Data Science — *Python, MATLAB*
Topic: *"Last-Mile Delivery Optimization with Recurrent Neural Network"*

- Proposed and built a pair-wise Recurrent Neural Network with a customized attention-based mechanism to predict the deviated path a driver would follow, adaptive to circumstances, compared to the theoretically planned shortest path.
- Designed an iterative Sequence Generation Algorithm used after model training to identify the first stop of a route that yields the optimal operational cost and achieves the global efficiency of all routes under 120 different simulations.
- Improved prediction accuracy from this model **by 15%** compared to the LSTM encoder-decoder and pointer network.

**Independent Research - Harvey Mudd College** — January 2023 - May 2023
Topic: *"Explore Constraints on Unitary Recurrent Neural Networks"* (**paper**) — *Python*

- Proved that bounding weight matrices within the unitary group exhibits superior parameters training, can be leveraged by stacking multiple recurrent layers, and thus can prevent the RNN vanishing and exploding gradient problems.
- Implemented a Unitary Recurrent Neural Network from scratch, adapting Arjovsky's paper while introducing a novel update rule constraint from prior proof to demonstrate the new model's practicality across Language Modeling tasks.

**Climate Research Lab - Harvey Mudd College** — August 2022 - December 2022
Research Assistant — *Python*

- Built an end-to-end Machine Learning pipeline from scratch to estimate PM2.5 particle size and sea level rise.
- Proposed a hybrid CNN-LSTM model which improved previous research's accuracy by 6% with only 4 hidden layers.

## SOFTWARE ENGINEERING EXPERIENCE

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

- Delivered a topology-preserving Dimension Reduction algorithm for Meta AI's CommerceMM model, achieving a **93% reduction in parameters** and a 2% improvement in accuracy for labeling **100,000,000+ Marketplace images**.
- Proposed a model-agnostic Window Selection algorithm partially addressed Vision Transformer's exponential scaling issue and helped researchers conduct performance analysis of the model and compare that 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, MongoDB, gRPC, Kubernetes, Terraform, AWS*

- 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.

## PROJECTS

**LendingClub Economic Risk Assessment Large Scale Data Analysis** — *R, tidyverse, ggplot*

- 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.
- Designed a k-NN model using my proposed Mahalanobis distance metric to forecast interest rates with 94% accuracy.