Hoang Viet Chu

1050 North Mills Avenue, Claremont, CA | hoachu@students.pitzer.edu | 909-407-6637 | github.com/HoangChu-Claremont

EDUCATION

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

- **GPA**: 3.7 / 4.0 (Major GPA: 3.9 -- Math and CS classes cross-registered at **Harvey Mudd College**)
- **Classes**: Machine Learning, Mathematics of Topic Modeling, Neural Network, Computer Systems, Linear Models.
- **Teaching Assistant -- 1 year**: Data Structures and Algorithms (in C++), Intermediate Probability, Abstract Algebra.
- **Awards:** Silver Medal: Asian Pacific Mathematical Olympiad, 1st Prize: National Mathematical Olympiad, USACO Platinum, 1st Prize: Citadel Datathon, Round 3: Google Code Jam, Winner: SIG Challenge, Codeforces Master.

RESEARCH EXPERIENCE

Periwinkle Trading August 2023 - Present
Quantitative Researcher - Contract *Python*

- Developed an end-to-end Generalized Regression model with Elastic Net regularization, achieving a **R^2 ~ 0.55** and satisfying all Gauss-Markov assumptions with Weighted Correlation Matrix and Variance-Stabilizing Transformation.
- **Reduced 30% runtime** of the model pipeline by **avoiding matrix inversion**: removed correlated features with PCA, optimized Beta(s) with QR Factorization by Givens Rotations, and estimated Covariance Matrix with Factor Models.

University of Southern California June 2023 - August 2023
Undergraduate Research Intern *Python, MATLAB*
Topic: *"Last-mile Delivery Optimization with Recurrent Neural Network"*

- Proposed a Pair-wise Attention-based Neural Network to predict and suggest stop sequences having lower total driving time than the theoretical path from Traveling Salesman algorithm under additional constraints e.g weather.
- Designed a new iterative sequence generation algorithm that identified the first stop of a route yields the lowest operational cost in model training and solved the problem with unordered input sequence from seq2seq model.
- Built from scratch a Transformer-based Deep Learning model and optimized the runtime with Adam SGD optimizer.

Independent Research January 2022 - May 2022
Topic: *"Explore Constraints on Unitary Recurrent Neural Networks"* (**paper**) *Python*

- Proved that the norm of Unitary Matrices are bounded, thus they can prevent RNN vanishing and exploding gradients.
- Exploited the orthogonality of Hermitian matrices to create **a new update rule** that **ensures the next weight matrix is always a unitary matrix** (complex spectral theorem) and **is computed efficiently** by eigen-decomposition.
- Implemented **from scratch** a Unitary Recurrent Neural Network with the above new update rule, adapting Arjovsky's paper, to demonstrate that the new model doesn't explode or vanish and can be used for LLM and Topic Modeling.

Citadel Datathon Winner - LendingClub Risk Assessment (26,000,000 rows) *Python*

- Discovered from hypothesis tests a negative correlation of income and (interest rate + risk subgrade), thus deployed a kNN model and found out LendingClub assigned higher-than-expected risk subgrades to similarly risky borrowers.
- Ideated and built a customized Mahalanobis distance metric for K-Means algorithm, allowing for the final discovery that Lending Club failed to detect previously defaulted borrowers who changed their information.

SOFTWARE ENGINEERING EXPERIENCE

Meta May 2022 - August 2022
Engineering Intern (received return offer) *Python, Hack*

- Proposed adaptation of two Microsoft Research papers on Vision Transformer (ViT) to Meta's ResNet50 model and customized a Window Selection algorithm addressing the ViT's scaling issue with **100,000,000+** Marketplace images.
- Independently built a prototype of the ViT model resulting in a **2% accuracy improvement** on 2,000,000 images, and initiated to implement and train Contrastive Captioners model, helping the team test new text generative algorithms.

Cohost.ai (cohost.ai) June 2021 - August 2021
Engineering Intern *C++*

- Ideated a full-stack design and built from scratch an NLU-based conversational chatbot that serves 2000+ daily users.
- Collaborated with full-time engineers to build an IPC message queue to avoid message loss if chatbot crashes.
- Designed Tree Serialization algorithms reducing company web PageSpeed from 30 secs to a consistent 0.8 second.