Hoang Chu

chu.hoang322@gmail.com | 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**: 3.68 / 4.0
- **Coursework**: Data Structures and Algorithms (Teaching Assistant), Intermediate Probability and Linear Algebra, Statistical Linear Models, Machine Learning (MIT OCW), Mathematics of Topic Modeling, Abstract Algebra.

AWARDS

- Winner - Citadel Invitational Datathon
- Winner - SIG Coding Challenge
- Full-ride Trustee Scholarship
- USACO Gold

RESEARCH EXPERIENCE

University of Southern California Los Angeles, CA
Undergraduate Research Summer Scholar in Operations Management and Data Science May 2023 - August 2023

Mathematics Research - Explore Unitary Recurrent Neural Networks (uRNN) Claremont, CA
Independent Research in Abstract Algebra January 2022 - May 2022

- Proved that restricting weight matrices to be within the unitary group can prevent vanishing and exploding gradients.
- Customized a sample Unitary Recurrent Neural Network model, which adapts Arjovsky et al. work but having a self-designed updating rule, to demonstrate the practicality of the previous proof in language modeling tasks.

WORK EXPERIENCE

Meta Menlo Park, CA
Software Engineer Intern May 2022 - August 2022

- Built Deep Learning models based on Transformer architecture, which helped improve the accuracy of the team's various Ads and Commerce classification tasks by 2% from 100 million images compared to baseline ResNet50.
- Researched the integration of Microsoft Research's Swin Transformer V2 paper into FAIR's CommerceMM model.
- Implemented and pre-trained Contrastive Captioners model enabling the team to test new text generative approaches.
- Wrote Dataswarm pipelines preparing the pre-training dataset and downstream tasks using internal operators.

CoHost.ai Hanoi, Vietnam
Software Engineer Intern June 2021 - August 2021

- Developed a chatbot that helps ~200+ daily users choosing suitable houses for rent.
- Built an Intent-Entity label matrix after finding connections from an eigenvalue to Markov chain's equilibrium point.
- Constructed a statistical language model that labels Intent-Entity from users' input texts, achieving a 95% accuracy.

PROGRAMMING PROJECTS

Risk Assessment Large Scale Data Analysis, Team of 4 (source)

- Detected from a 26 million dataset that given similar risk profiles, LendingClub charges higher interest rates than lending counterparts, and when previously similar risk profiles had multiple defaults, it fails to adjust interest rates.
- Self-designed a distance metric that helps the team's k-NN algorithm only need a small subset of features to classify groups of borrowers by interest rate and achieve a 94% accuracy in predicting interest rate of a new borrower.

Optimized Course Searching Website, Independent | Python (source)

- Designed a non-clustered database indexing with self-designed tree data structures and an accurate course-equivalent matching algorithm that helped reduce the search engine's response time from 25 to consistently 0.8 seconds.

TECHNICAL SKILLS

Languages: Python, C++, Java, JavaScript, SQL
Frameworks and Tools: Pandas, Numpy, Scikit-learn, Flask, MongoDB, Docker