Henry (Hoang) Chu

1050 North Mills Avenue, Claremont, CA | hchu88347@gmail.com | 909-381-6851 | github.com/Renaxis-Math

EDUCATION

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

- GPA: 3.7 / 4.0 (Math and CS classes cross-registered at Harvey Mudd College).
- **Classes**: Data Structures and Algorithms (**TA)**, Advanced Probability (**TA**), Abstract Algebra (**TA)**, Combinatorics, Statistics, Machine Learning, Neural Network, Natural Language Processing, Statistical Linear Models, Mathematics of Data Science, Computational Theory, Advanced Linear Algebra, Programming Languages, Computer Systems.

AWARDS

- 9 / 82 teams - Southern California ICPC Contest
- Gold Medal (9 / 382): National Math Olympiad
- 3rd Ranked: National Math Modeling Olympiad
- Silver Medal: Asian Physics Olympiad
- Codeforces Master (2147 rating - Top 3%)
- Top 1% (Round 3) Google Code Jam 2022
- 4th Place - Citadel Terminal Algorithm Contest
- 1st Place - Citadel Summer Invitational Datathon

WORK EXPERIENCE

Periwinkle Trading (periwinkletrading.com) August 2023 - Present
Quantitative Researcher - Contract (report to the CEO) *Technologies:* ***Pandas, Numpy, Scikit-learn, Apache Spark***

- Ideated using a Support Vector Regression to predict 1-day Equities prices and developed a custom NLP-based feature selection from **20GB of alternative data** to generate trading signals, **improving the Sharpe ratio from 1.3 to 2.1**.
- **Improved** weighted correlation coefficient **from .64 to .81** with **45% reduced features** by utilizing geometry, linear algebra, and binary search algorithm to find LASSO penalty term that max(VIF score) < 3 and minimizes AIC.
- **Scaled** a price/yield calculation engine in C++ that handled increased **from 30000 to 65,000** daily requests, and created an integration testing suite in C++ for all types of trades and also implemented metrics logging to Datadog.
- **Reduced** model **computational complexity from O(n) to O(log n)** by utilizing Double Heaps and Binary Search.

University of Southern California June 2023 - August 2023
Undergraduate Research Intern *Technologies:* ***Pandas, Numpy, Scikit-learn, OR-Tools***
Topic: *"Solving the Vehicle Routing Problem with Recurrent Neural Network (RNN)"*

- **Tackled neglection of real-time factors** (e.g traffic jam) in the Traveling Salesman shortest-Path (TSP) algorithm by proposing and developing from scratch a customized attention-based Recurrent Neural Network model in C++.
- **Solved seq2seq's adaptability issue** with unordered sequences by convincing the team to use LSTM encoder-decoder (capture global instead of local information) with **a model built in Python** that **saved 15% time** on average than TSP.
- Proved convergence to suboptimal loss of incremental gradient method techniques in network optimization, and **utilized past research** in Group Theory to avoid RNN's vanishing and exploding gradients in model deployment.

Meta May 2022 - August 2022
Engineering Intern *Technologies:* ***PyTorch, Pandas, Numpy, Scikit-learn, Hadoop (MapReduce), kDB***

- Researched Vision Transformer (ViT) papers and delivered in Python a working model to replace Meta's ResNet50
- Resolved ViT's Time Limit Issue in hyperparameter tuning by designing, discussing, and implementing independently a new Window Selection Optimization algorithm that **helped Vision Transformers compute 30% faster**.
- **Improved label accuracy by 2%** on **100,000,000+** Marketplace training data after researched, pre-trained in Python, and proposed Google's Contrastive Captioners encoder-decoder (arxiv.org/abs/2205.01917) in lieu of Meta's CLIP.

PROJECTS

Data Science Project - *"Car Price Prediction on a 3 Million Dataset"* *R*

- Reduced 2/3 features while improving R^2 and satisfied all linear regression assumptions by performing a bounded optimal penalty value for Elastic Net Regression, utilizing VIF tracing with eigendecomposition and BIC scores.

Independent Research in Mathematics
Topic: *"Explore Constraints on Unitary Recurrent Neural Networks (uRNN)"* (**paper**) *Python*

- Proved that the norm of Unitary Matrices are bounded, thus they can prevent RNN vanishing and exploding gradients.
- Leveraged Hermitian matrices' orthogonality properties to **present a new gradient update rule**, ensuring subsequent weight matrices being unitary by complex spectral theorem and efficiently computed by eigen-decomposition.