Vinh-Hung (Tommy) Ly *Curriculum Vitae*

https://tommyly.net   tommy.ly@columbia.edu   me@tommyly.net

3041 Broadway, New York, NY 10027   +1 (718) 637-9597

Education

Columbia University, New York   May 2025
  Bachelor of Arts in Computer Science, Minor in Statistics   **GPA: 3.8/4.0**
  Selected Coursework: Statistical Analysis of Neural Data (PhD), Probabilistic Models and Machine Learning (PhD), Bayesian Statistics, Machine Learning, Analysis of Algorithms, Introduction to AI
  Planned: Elementary Stochastic Processes, Time Series Analysis, Game Theory

Royal Melbourne Institute of Technology Vietnam   November 2013
  Bachelor of Commerce

Honors & Awards

Cohen Scholar - Awarded a scholarship ($25,000) by the Steven & Alexandra Cohen Foundation, in recognition of perseverance and dedication to completing educational goals
Columbia School of General Studies Honor Society - One of 76 students inducted. Recognized for exceptional academic achievements, minimum GPA requirement of 3.8
Dean's List (4 semesters) - Recognized for outstanding academic performance with a GPA above 3.6

Research Experience

The Mortimer Zuckerman Mind Brain Behavior Institute, New York   Summer 2024 - Present
Research Assistant, Computational Neuroscience Lab   *Advisor: Prof. Liam Paninski*

- Developed semi-supervised pose estimation pipeline using Ensemble Kalman Smoother (EKS) and iterative self-training, achieving 10% improvement over baseline ensemble models
- Developed novel data augmentation methods for multi-session neural decoding using adaptive dropout based on neuron centrality measures and graph contrastive learning
- Manuscript in preparation: "Improving Pose Estimation through Ensemble Kalman Smoothing and Transfer Learning"

Columbia Irving Medical Center, New York   Summer 2023
Research Intern, Program for Mathematical Genomics   *Advisor: Prof. Raúl Rabadán*

- Developed hybrid Convolutional Neural Network-Transformer architecture for predicting chromatin accessibility from genomic sequences
- Implemented processing pipeline for 115 ATAC-seq datasets, achieving Pearson correlation of 0.5 through residual connections and batch optimization

New York Genome Center, New York   Summer 2022
Research Intern, Technology Innovation Lab   *Advisor: Prof. Sanja Vicković*

- Modified PySeq2500 package to repurpose Illumina HiSeq 2500 as automated fluorescence microscope, implementing background correction and focal plane detection using dask and xarray

Industry Experience

Uber Technologies Inc
Senior Regional Operations Manager, New York   2018 - 2020

- Designed spatial-temporal analysis framework using SQL to identify optimal cross-dispatch opportunities, analyzing over 500,000 trips to quantify market-specific patterns and driver utilization rates
- Implemented city-wide switchback experiments to test courier boosting thresholds, resulting in reduced time arrivals but revealing negative impact on dedicated courier metrics (trips/hour, supply hours); findings drove transition to granular trip-level optimization using Cobb-Douglas production function
- Built complex SQL pipeline (300 lines) to calculate Long-term Equitable Dispatch Incentive burn rates for cross-platform drivers across US/Canada, validating offline analysis against experimental results

Regional Operations Manager, Singapore   2017 - 2018

- Led Asia-Pacific marketplace optimization, developing analytics tools for surge pricing and ETA anomaly detection; implemented dispatch-time-limit surge mechanism improving marketplace efficiency across multiple cities
- Designed and executed regional experimentation program, generating highest regional experiment proposal rate (6+ in 4 weeks) and building automated analysis framework for marketplace dynamics

Operations Manager, Vietnam   2016 - 2017

- Led region's analytics initiatives including successful referral program and driver incentive optimization, executing 7 consecutive A/B tests with novel user segmentation that were replicated across Southeast Asia

Projects & Software

attract-repel embedding: Novel graph embedding technique for non-transitive relationships in networks
brainsets: A neural decoder framework with adaptive augmentation strategies for multi-session recordings
pseudo-labeler: Semi-supervised learning pipeline for neural data processing
atac-rna processing: Processing pipeline for chromatin accessibility and gene expression analysis
pyseq2500: Extended package for automated fluorescence microscopy with parallel image processing

Skills

Programming Languages	Python, Java, C
ML Frameworks	PyTorch, scikit-learn, Lightning AI, gcloud, aws, PostgreSQL

Activities

AI@Columbia: Co-founder. Built 600+ member community of faculty, researchers, and students. Organized industry speaker series and networking events sponsored by venture capital
Triathlon: Completed Ironman Maryland (2.4mi swim, 112mi bike, 26.2mi run) in 14.5 hours