Vinh-Hung (Tommy) Ly

tommyly.net tommy.ly@columbia.edu +1 (718) 637-9597

Education

Columbia University, New York May 2025
B.A. Computer Science and Statistics **GPA: 3.82/4.0**
Selected Coursework: Artificial Intelligence, Probabilistic Machine Learning (PhD), Statistical Analysis of Neural Data (PhD), Bayesian Statistics, Algorithms, Stochastic Processes, Time Series
Honors: Cohen Scholar ($25k), Columbia GS Honor Society (top 76 students), Dean's List (6 semesters)

Royal Melbourne Institute of Technology Vietnam - B. Com 2013

Research Experience

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

- Developed a semi-supervised computer vision pipeline (github: pseudo-labeler) for pose estimation by integrating an Ensemble Kalman Smoother with iterative self-training, achieving a 10% improvement over baseline ensemble models; manuscript in preparation.
- Designed novel data augmentation strategies for the github: brainsets project, implementing adaptive neural dropout (gist excerpt) informed by neuron centrality, thereby enhancing model robustness.

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

- Developed a hybrid CNN-Transformer deep learning model (github: atac-processing) for chromatin accessibility prediction, processing 115 ATAC-seq datasets, achieving a Pearson correlation of 0.5.

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

- Modified the PySeq2500 by implementing background correction and focal plane detection using dask and xarray

Projects

- **causal-bias-agent (early-stage)**: developing an agent-based AI pipeline for collider bias detection in the scientific literature, integrating DAG-based causal inference, retrieval-augmented generation, semantic embeddings, and prompt chaining.
- **attract-repel embedding article**: re-implemented pseudo-euclidean attract-repel embeddings for undirected graphs by assigning to each node distinct attract and repel vectors.

Industry Experience

Uber Technologies Inc, New York, Singapore & Vietnam 2016 - 2020
Senior Regional Operations Manager

- Designed and implemented a spatial-temporal analysis framework using SQL and Python to identify optimal cross-dispatch opportunities, analyzing 500k+ trips and uncovering market-specific patterns for improved driver utilization.
- Built complex SQL pipelines to calculate incentive burn rates across US/Canada, validating offline analyses with experimental results for enhanced data reliability.

Skills & Activities

Technical Skills: Python, Java, C, SQL, PyTorch, scikit-learn, Lightning AI, Google Cloud Platform, AWS, PostgreSQL, Git, Linux

Activities: Co-founder of AI@Columbia (600+ member community), Completed Ironman Maryland (14.5 hours)