

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 for pose estimation by integrating an Ensemble Kalman Smoother with iterative self-training, achieving a 10% improvement over baseline ensemble models; manuscript in preparation ([github: pseudo-labeler](#)).
- Implemented adaptive neural dropout ([gist excerpt](#)) informed by neuron centrality, thus improving the robustness of the foundation model for neural decoding ([github: brainsets](#)).

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 for chromatin accessibility prediction, processing 115 ATAC-seq datasets, achieving a Pearson correlation of 0.5 ([github: atac-processing](#)).

New York Genome Center, New York

Summer 2022

Research Intern, Tech Innovation Lab

Advisor: Prof. Sanja Vicković

- Created background correction and focal plane detection using dask and xarray for PySeq2500 microscopy

Projects & Open Source Contributions

- **attract-repel embeddings:** Implemented pseudo-Euclidean embeddings for link prediction in graphs; contributed core algorithm to PyTorch Geometric, the leading graph deep learning library ([pull request link](#)).
- **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.
- **talent-nlp:** Implemented a modular NLP pipeline with FastAPI, spaCy, and transformer models for document segmentation; deployed as a RESTful API with async processing ([github: o1a-assessment](#)).

Industry Experience

Uber Technologies Inc, New York, Singapore & Vietnam

2016 - 2020

Senior Regional Operations Manager

- Architected spatial-temporal analysis frameworks and complex SQL pipelines for cross-dispatch optimization and incentive burn rate calculation, processing 500k+ trips and validating predictions with experimental results for enhanced driver utilization and data reliability.

Skills & Activities

Technical Skills: Python, Java, C, SQL, PyTorch, scikit-learn, spaCy, Transformer models, FastAPI, RESTful APIs, Lightning AI, GCP, AWS, PostgreSQL, Git, Linux

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