Thien Le

550 Memorial Drive, #16A1
Cambridge, MA 02139, USA

Email: thienle@mit.edu
Google Scholar: Thien Le

EDUCATION

Massachusetts Institute of Technology (MIT), EECS-CSAIL Jegelka Group　　　Sep 2019 - May 2024
Master of Science (SM) and Doctor of Philosophy (PhD) in EECS　　　GPA: 5.0/5.0
　　Master thesis: Training invariance and the low-rank phenomenon: beyond linear networks

University of Illinois at Urbana-Champaign (UIUC)　　　Sep 2016 - May 2019
Bachelor of Science in Mathematics & Computer Science　　　GPA: 3.93/4.00
　　Awarded 'Most Outstanding Undergraduate Major Award in Mathematics & CS'

RESEARCH INTERESTS

Geometric Deep Learning. (Deep) Learning Theory. Optimization. Graph Theory and Graph Limits

PROJECTS

Graduate researcher under Stefanie Jegelka

Learning with complex (non-manifold) biological structure　　　Sep 2019 - Ongoing
- Studying the algebraic geometry and combinatorics of tree space (a non-manifold moduli space)
- Implemented a new algorithm in Haskell for tree inference based on continuous optimization of that space
- Gave a seminar talk at VinAI research day

Implicit regularization of deep ReLU neural network　　　Sep 2020 – Oct 2021
- Generalizing theoretical results on margin and low rank bias in linear network to deep ReLU architectures
- Established a family of invariants during late-stage training of said architectures under gradient flow/descent
- Publish in ICLR2022 (Spotlight)

Transferability of graph neural network via graph limits (graphops)　　　Oct 2022 - Feb 2023
- Study limits, approximation and transferability of finite graph neural networks via contemporary theory of graph limit (graphop - Backhausz and Szegedy)
- First transferability result for sparse graph sequences (small molecules, grid graphs for images, hypercubes, etc.)
- Publish in NeurIPS2023

Learning theory for invariant/equivariant concept classes　　　Feb 2023 - Ongoing
- Study learning theory of deep neural network architectures designed to be invariant/equivariant under certain group actions such as CNN, DeepSets, SignNets, etc
- Show lower-bound for learning simple graph neural nets and frame-averaging of simple neural nets under correlational statistical queries (CSQ) models
- Publish in ICLR2024 (Spotlight)
- Extending to staircase-complexity models

Graphon signal sampling　　　Oct2022 - Jul2023
- Develop a sparse scheme of node sampling for graph signals under graph limit models (graphons) akin to Shannon-Nyquist theorem
- Prove a Poincare inequality and consistency of such sampling scheme, demonstrate empirical performance on real-world datasets
- Connect the geometry of mixture of graphons to spectral clustering under mixture models
- Publish in ICLR2024 (Spotlight)

Statistical Phylogeny Gene Tree Estimation　　　Jun 2018 – Ongoing
- Implementing a scalable phylogeny estimation algorithm with strong statistical guarantees
- Implemented parallelism and algorithmic optimization (e.g. LCA in $O(1)$ query and $O(n)$ preprocessing)
- Maintaining 5000 lines of C code at constraint_inc, tested with BlueWaters supercomputer
- Three papers published in conference and journal

TEACHING AND INTEREST

- Teaching: Fundamentals of Statistics (Summer 2023) and Deep Learning (Fall 2023) @ MIT-EECS; both core algorithm classes @ UIUC-CS
- Reviewer for ICLR, AISTATS, ICML, NeurIPS, ECML-KPDD.
- Graduate math courses @MIT: Stochastic Calc., Geom. of Manifolds, AG I, Intro to Representation Thr., Analysis of Boolean Functions