

Yilun Kuang

+19299776340 | yk2516@nyu.edu | github.com/YilunKuang | [linkedin.com](https://www.linkedin.com/in/yilunkuang)

EDUCATION

New York University

cumulative GPA: 3.85/4.00

B.S Honors Mathematics; B.A Computer Science

Sep. 2020 – May 2023

- Relevant coursework: Cloud and Machine Learning (Grad), NLP with Representation Learning (Grad), Math Tools for Data Science (Grad), Mathematics of Deep Learning (Grad), Foundations of Machine Learning (Grad), Honors Numerical Analysis, Honors Analysis I & II, Parallel Computing, Object-Oriented Programming

Pitzer College

cumulative GPA: 3.9/4.0

B.A Cognitive Science; B.A. Philosophy

Sep. 2019 – May 2020

- Relevant coursework: Intro to Cognitive Science, Foundations of Neuroscience, Topics in Neurophilosophy

RESEARCH EXPERIENCE

Flatiron Institute (Simons Foundations)

Jan 2022 – Present

Research Intern (Advisor: Prof. SueYeon Chung)

New York, NY

- Developed a manifold compression based self-supervised learning algorithms. Distributed Training on ImageNet using 32 A100 GPUs gives us on-par performance compared to other state-of-the-art algorithms like SimCLR and Barlow-Twins. Preparing for ICLR 2023 Submission.

NYU Machine Learning for Language Group

June 2021 – Present

Natural Language Processing Research Assistant (Advisor: Prof. He He)

New York, NY

- Fine-tune BART and GPT2 on In-Distribution datasets (ID) and implement ROUGE, PPL, AUROC evaluation metrics on several Out-of-Distribution (OOD) text summarization datasets on NYU HPC Greene clusters.
- Run experiments on the Singularity Containers and use the Slurm resource allocation system for large-scale model pretraining, fine-tuning, and evaluations. Monitor the hardware and model performance with customized NVIDIA CLI interface. Hands-on model checkpoints, loggings, and debugging.
- Follow up on literature on out-of-distribution detection, prompt tuning, adapters, and develop a new algorithm for unsupervised domain adaptation in text summarization.

PROJECT EXPERIENCE

SiFT Algorithm for Improved Generalization

Feb 2021 – May 2021

- Implemented Scale-invariant-Fine-Tuning (SiFT) algorithm and the Smoothness-Inducing Adversarial Regularization and Bergman Proximal Point Optimization (SMART) algorithm for virtual adversarial training on BERT and DeBERTa and achieved better adversarial robustness and model generalization.
- Implemented a novel model generalization benchmark: Perform fine-tuning for BERT and DeBERTa on In-Domain Twitter Hate Speech dataset and fine-tune on 10 percentage of Out-of-Domain UCI Sentiment dataset and evaluate the testing result.

NTK and Mean Field Limit for Overparametrized Network Training Dynamics

Feb 2022 – May 2022

- Course Project in CSCI-GA 3033 Mathematics of Deep Learning. Compared the lazy learning (NTK) vs. active learning (Mean Field Limit) regimes for overparametrized neural networks. Project Report Link
- Numerically simulated two regimes for finite neural networks and compare the training dynamics.

Poisson-Gamma Neural Variability in the Visual Cortex

Oct 2020 – Dec 2020

- Derived a poisson-gamma probabilistic model to characterize the noise distribution in visual neuron responses.
- Simulated neural population encoding and performed maximum likelihood decoding and statistical tests to validate the poisson-gamma model in Matlab

SKILLS/OTHER

- Honors and Award: Nicholas and Andrea Ferrara Research Scholar, Dean's List, DURF Grant Recipient
- Competition: Meritorious Winner in 2021 Mathematical Contest in Modeling (MCM) (Top 7%)
- Language: C++, C, Python, Java, Matlab, Julia, SQL, Bash
- Framework: PyTorch, TensorFlow, Transformers, Scikit-Learn, CUDA, OpenMP, MPI