Qiaobo Li

Department of Computer Science, University of Illinois at Urbana-Champaign
+1 4479021225 | qiaobol2@illinois.edu
Graduation year: May 2027

EDUCATION

PEKING UNIVERSITY 09/2018–06/2022
School of Mathematical Sciences

- Undergraduate, Major in Statistics and Probability, Overall GPA: **3.79**/4; Major GPA:**3.80**/4
- Programming & Software: C, C++, Python, MATLAB, R; Latex
- English Proficiency: TOEFL 110, GRE Verbal 157, Quantitative 170, Writing 3.5

Awards and Honors

Academic Excellence Award, Peking University	11/2020&11/2021
Annual Scholarship, Peking University	11/2020&11/2021
3rd Prize in Chinese Mathematical Olympiad	01/2018
2nd Prize in Chinese Mathematics Competition	11/2019

University of Illinois at Urbana-Champaign 08/2022–present
Department of Computer Science

- Ph.D., Major in Computer Science, Overall GPA (till now): 3.95/4
- Programming & Software: C, C++, Python, Pytorch, MATLAB, R; Latex
- Research interests: Federated Learning, Trustworthy Machine Learning, Deep Learning Theory, High-dimensional statistics.

Main Course and Score

Deep Learning Theory	4/4	Trustworthy ML	4/4
Adv Topics in Sec, Priv and ML	4/4	Adv Topics in NLP	4/4

RESEARCH EXPERIENCE

Understanding learning/generalization landscape for GD on multi-task/single-task learning 07/2021-08/2023
Supervised by Prof. Quanquan Gu, Department of Computer Science, UCLA

Trustworthy Machine Learning 02/2023-08/2023
Supervised by Prof. Han Zhao, Department of Computer Science, UIUC

Geometric Structures and Others in Machine Learning 09/2023-present
Supervised by Prof. Arindam Banerjee, Department of Computer Science, UIUC

WORK EXPERIENCE

Machine Learning Engineer Intern 05/2025-08/2025
Responsible Recommendation System, TikTok Inc.

Key Contributions:

- Select state-of-the-art vision–language models and appropriate fine-tuning methods and aim to achieve acceptable accuracy on the given datasets.
- Explore strategies for imbalanced datasets, such as combining class-specific fine-tuned models and rebalancing the dataset through resampling or reweighting.
- Compare and analyze how different frame-extraction strategies affect video classification performance of vision–language models.
- Design and study a router that dispatches video inputs between larger and smaller vision–language models to balance accuracy and computational cost.

Recognition: Exceeds Expectations (E) performance rating.

PUBLICATIONS

- **Sketched Gaussian Mechanism for Private Federated Learning**
  Qiaobo Li, Zhijie Chen, Arindam Banerjee
  NeurIPS Spotlight, 2025

- **Sketched Adaptive Federated Deep Learning: A Sharp Convergence Analysis**
  Zhijie Chen, Qiaobo Li, Arindam Banerjee
  NeurIPS Poster, 2025
- **On the Power of Multitask Representation Learning with Gradient Descent**
  Qiaobo Li, Zixiang Chen, Yihe Deng, Yiwen Kou, Yuan Cao, Quanquan Gu
  AISTATS Poster, 2025
- **Loss Gradient Gaussian Width based Generalization and Optimization Guarantees**
  Arindam Banerjee, Qiaobo Li, Yingxue Zhou
  AISTATS Oral, 2025
- **Differentially Private Post-Processing for Fair Regression**
  Ruicheng Xian, Qiaobo Li, Gautam Kamath, Han Zhao
  ICML Poster, 2024
- **Sketching for Distributed Deep Learning: A Sharper Analysis**
  Mayank Shrivastava, Berivan Isik, Qiaobo Li, Sanmi Koyejo, Arindam Banerjee
  NeurIPS Poster, 2024
- **Sketched Gaussian Mechanism on Matrix for Private Federated LoRA**
  Qiaobo Li, Zhijie Chen, Arindam Banerjee
  Openreview, 2025
- **Beyond Johnson-Lindenstrauss: Uniform Bounds for Sketched Bilinear Forms**
  Rohan Deb, Qiaobo Li, Mayank Shrivastava, Arindam Banerjee
  Openreview, 2025

Coding Projects

- **Federated Learning with Communication-Efficient and Private Algorithms**
  - **Objective**: Advance federated learning by developing a novel algorithm and benchmarking it against state-of-the-art communication-efficient and differentially private methods for both vision and language tasks.
  - **Key Contributions**:
    - Designed and implemented different communication-efficient and private federated learning algorithms, leveraging sketching techniques to enhance communication efficiency.
    - Conducted rigorous evaluations of the algorithm alongside **FetchSGD** and **1bit-ADAM**, benchmarking performance on vision and language tasks:
      - Trained **ResNet101** and fine-tuned **ViT-based model** on **CIFAR-10** for vision tasks.
      - Fine-tuned **BERT** on **SST2** from the GLUE benchmark for sentiment analysis.
    - Investigated spectral properties of optimization dynamics to gain insights into model optimization behavior during training and fine-tuning by analyzing **loss Hessian** and **predictor Hessian spectrum structures** using the **SLQ algorithm**, revealing trends in curvature and stability.
    - Demonstrated improvements in **accuracy**, **convergence speed**, and **resource utilization**, while maintaining competitive communication overhead across diverse server optimizers, including **SGD**, **SGD with momentum**, **ADAM**, and **AdaClip**
- **Fair and Private Post-Processing for Regressors**
  - **Objective**: Developed and conducted a fair and private post-processing algorithm to process regressors on the **Communities & Crime** and **Law School** datasets.
  - **Key Contributions**:
    - Investigated and quantified trade-offs among **test error**, **fairness violation**, and **differential privacy** by implementing **differential privacy mechanisms** and fairness-aware evaluation metrics.
    - Enhanced model performance while balancing **privacy guarantees** and **fairness constraints** through systematic analysis of the interplay between privacy loss (ε), fairness measures, and predictive accuracy.
- **Loss Gradient Gaussian Width: Generalization and Optimization Guarantees**
  - **Objective**: Analyze generalization and optimization trends during training by investigating gradient and featurizer properties across models and datasets.
  - **Key Contributions**:
    - Trained **ResNet18**, **Feedforward Networks (FFN)**, and **CNN** on **CIFAR-10**, **MNIST**, and **Fashion-MNIST** using **SGD**, **ADAM**, and **ADAMW**.
    - Conducted comprehensive gradient analyses, identifying trends in **Gradient Domination Ratio**, **L2 norm of gradients**, and **L0/L1 norms of featurizers**, and examined gradient coordinate structures to understand their relationship to generalization performance.
    - Assessed the interplay between model architectures, optimizers, and dataset characteristics, providing insights into how these factors influence optimization dynamics and generalization behavior.