BOWEN LEI

2250 Dartmouth St Apt 626, College Station, Texas, 77840, U.S.A.
979 446 1316 ⋄ bowenlei@stat.tamu.edu ⋄ https://stevenboys.github.io

RESEARCH INTEREST

I am interested in **efficient & reliable** deep learning and **Bayesian** machine learning. My current research is investigating how to improve the reliability & efficiency of deep learning systems to achieve Pareto optimality between decision reliability & computing resources and model performance.

· **Efficient Training & Inference**: Sparse Training, Pruning, Data Distillation, Few-shot Learning
· **Reliable Deep Learning**: Uncertainty Quantificatio, Robust Generalization, Adaptation
· Domains: Computater Vision, Natural Language Processing, Recommendation Systems, Sciences

EDUCATION

Texas A&M University, Ph.D. in Statistics — *Aug 2019 - May 2024 (Expected)*
Advisor: Prof. Bani K. Mallick — GPA: 3.89/4

Renmin University of China, B.S. in Statistics — *Sep 2015 - Jun 2019*
Graduated as the 2nd in 35 — GPA: 3.87/4
Minors in Mathematics and Economic Statistics

Yale University, Visiting student in Biostatistics department — *Oct - Dec 2018*

PROFESSIONAL EXPERIENCE

Video Engineering team, Apple — *May – Aug 2022*
ML Researcher Intern (Manager: Andrew Bai; Fang-Yu Lin) — Cupertino, CA, USA

· Project: image quality assessment & improvement and image denoising pipeline.

Search team, JD.COM American. — *May – Jul 2021*
Research Intern (Manager: Yun Xiao; Xi Xiong) — Mountain View, CA, USA

· Project: cold start problems in recommendation system using cross-domain and group information.

Department of Data Modeling, Percent. — *Feb – Apr 2019*
Modeling Analyst Spring Intern (Manager: Taiyun Wei) — Beijing, China

· Researched on attention architecture in machine translation and natural language processing.
· Project: building machine translation system that can adapt to multiple domains and languages.

PUBLICATIONS

[1] **Bowen Lei**, Ruqi Zhang, Dongkuan Xu, Bani K. Mallick, "Improving Confidence Calibration and Reliability in Sparse Training", submitted in 2022.

[2] **Bowen Lei**, Dongkuan Xu, Ruqi Zhang, Shuren He, Bani K. Mallick, "Accelerating and Stabilizing Sparse Training", submitted in 2022.

[3] **Bowen Lei**, Tanner Quinn Kirk, Anirban Bhattacharya, Debdeep Pati, Xiaoning Qian, Raymundo Arroyave, Bani K. Mallick, "Bayesian Optimization with Adaptive Surrogate Models for Automated Experimental Design", **Nature** Computational Materials 7, 194 (2021).

[4] Lee, Se Yoon, **Bowen Lei**, and Bani K. Mallick. "Estimation of COVID-19 spread curves integrating global data and borrowing information." **PloS one** 15.7 (2020): e0236860.

[5] Shaoyi Huang, **Bowen Lei**, Dongkuan Xu, Hongwu Peng, Mimi Xie, Caiwen Ding, "Fast Dynamic Sparse Training via Better Exploration-Exploitation", submitted in 2022.

[6] Lei Zhang, Jie Zhang, **Bowen Lei**, Subhabrata Mukherjee, Xiang Pan, Bo Zhao, Caiwen Ding, Yao Li, Dongkuan Xu, "Efficient Gradient-matching Data Distillation", submitted in 2022.

[7] Yue Xiang, Dongyao Zhu, **Bowen Lei**, Dongkuan Xu, Ruqi Zhang, "Efficient Sampler for Discrete Distributions", submitted in 2022.

[8] Shaoyi Huang, Haowen Fang, Kaleel Mahmood, **Bowen Lei**, Nuo Xu, Bin Lei, Yue Sun, Dongkuan Xu, Wujie Wen and Caiwen Ding, "Sparse Training in Spiking Neural Network", submitted in 2022.

RESEARCH EXPERIENCE: EFFICIENT DEEP LEARNING

Accelerating Sparse Neural Network Training. *Sep 2021 – Oct 2022*

· Leader: Bowen Lei. Mentor: Bani K. Mallick, Dongkuan Xu, Ruqi Zhang.
· Researched on accelerating sparse robust training with adaptive variance reduction.
· Submitted in 2022.

Dynamic Sparse Training for Diffusion Model. *Oct 2022 – Present*

· Leader: Bowen Lei. Mentor: Bani K. Mallick.
· Researched on dynamic sparse training in diffusion model to achieve training and inference efficiency.
· Preparing a conference paper in 2023.

Efficient Gradient-matching Data Distillation. *Jul 2022 – Nov 2022*

· Mentor: Dongkuan Xu. Collaborator: Lei Zhang, Jie Zhang.
· Researched on efficient gradient-matching data distillation via pre-trained models.
· Submitted in 2022.

Sparse Training in Spiking Neural Network. *Sep 2022 – Nov 2022*

· Mentor: Caiwen Ding. Collaborator: Shaoyi Huang.
· Researched on sparse training in spiking neural network via the rigged lottery.
· Submitted in 2022.

RESEARCH EXPERIENCE: RELIABLE DEEP LEARNING

Reliable Sparse Neural Network Training. *Mar – Oct 2022*

· Leader: Bowen Lei. Mentor: Bani K. Mallick, Ruqi Zhang, Dongkuan Xu.
· Researched on sparse neural network with good calibration and comparable accuracy.
· Submitted in 2022.

Reliable and Efficient Out-of-distribution Detection. *Oct 2022 – Present*

· Leader: Bowen Lei. Mentor: Bani K. Mallick, Dongkuan Xu, Ruqi Zhang.
· Researched on improving reliability and efficiency of out-of-distribution detection.
· Preparing a conference paper in 2023.

Reliability in Data Distillation. *Nov 2022 – Present*

· Mentor: Dongkuan Xu. Collaborator: Dongyao Zhu, Jie Zhang.
· Researched on improving reliability of gradient-matching data distillation.
· Preparing a conference paper in 2023.

Local Calibration for Imbalanced Classification. *June 2022 – Present*

· Mentor: Dawei Zhou, Dongkuan Xu. Collaborator: Longfeng Wu.
· Researched on improving local calibration of minority class in imbalanced image classification.
· Preparing a conference paper in 2023.

RESEARCH EXPERIENCE: BAYESIAN MACHINE LEARNING

Image Classification with Physics-informed Features. *Nov 2020 – Present*

· Leader: Bowen Lei. Mentor: Bani K. Mallick, P.R. Kumar.
· Researched on variational autoencoder Model for image classification with Physics-informed Features.
· Preparing a journal paper in 2023.

Bayesian optimization with Adaptive Surrogate Models. *Aug 2019 – June 2021*

· Leader: Bowen Lei. Mentor: Bani K. Mallick, Anirban Bhattacharya, Debdeep Pati.
· Combined Bayesian optimization and adaptive surrogate models for autonomous experiment design.
· Accepted in npj Computational Materials - **Nature**.

Bayesian Hierarchical Richards Model in COVID-19. *Mar – July 2020*

· Mentor: Bani K. Mallick. Collaborator: Se Yoon Lee.
· Used Bayesian hierarchical Richards model in COVID-19 for trend prediction and intervention analysis.
· Accepted in **PLOS ONE** 2020 and developed R package **BHRM**.

PROFESSIONAL TALKS

Efficient and Reliable Sparse Training *Nov 2022*

· CSC 791&591: Advanced Topics in Efficient Deep Learning, NC State University.

Machine Learning in COVID-19 and Epidemiology *Apr 2021*

· STAT 21019063: Data Science in Action, Renmin University of China.

PROFESSIONAL SERVICES

ICML 2022: Conference Reviewer. Feb – Mar 2022
KDD 2022: Conference Reviewer. Mar – Apr 2022

TEACHING EXPERIENCE

Teaching Assistant, Texas A&M Univ., College Station, TX

· STAT 605, Advanced Statistical Computation (Prof. Pati) *Fall 2022*
· STAT 641, Methods of Statistics (Prof. Ghosh) *Spring 2022*
· STAT 657, Advanced Programming Using SAS (Prof. Kincheloe) *Spring 2021*
· STAT 645, Applied Biostatistics (Prof. Sinha) *Fall 2020*

HONORS AND AWARDS

Bachelor of Science (B.S.)

- First Prize, Student Scholarship for Excellent Academic Performance (top 2) *2017*
- Grand Prize in the 19th "Innovation Cup" Academic Research Competition (top 1%) *2017*
- First Prize, University Extracurricular Academic and Technology Competition (top 0.1%) *2017*
- First Prize, Fei Xiaotong Scholarship for Excellent Academic Performance (top 2) *2016*

EXTRACURRICULAR ACTIVITIES

- Assistant editor of Capital of Statistics *2017 - Present*
- Member and editor of CluBear *2017 - 2019*
- Organizer and volunteer of the China R Conference *2017 - 2019*
- Statistical consultant of Visualization and Visual Analytics Group in Peking Univ. *2018*
- Minister of Young Volunteers Association of the Department of Statistics *2016 - 2017*

SKILLS

Programming Python, R, SQL, C++, LaTeX
Tools TensorFlow, PyTorch, Keras, Scikit-learn