

# Dongbai Li

Third-year undergraduate student, Tsinghua University, Beijing, China  
ldb22@mails.tsinghua.edu.cn

## RESEARCH INTERESTS

---

Trustworthy AI, OOD generalization, Domain generalization, Data-centric AI

## EDUCATION

---

**Tsinghua University**, Beijing, China  
B.E. in Computer Science and Technology

9 2022 — Present  
Cumulative GPA: 3.87

## RESEARCH EXPERIENCE

---

**Prof. Huan Zhang's lab**

*Visiting research intern*

University of Illinois Urbana-Champaign (UIUC)

7 2024 — Present

- **Performance Prediction**

Subpopulation shifts, where the distribution of subgroups between training and target datasets differs, pose significant challenges to machine learning models, making performance prediction prior to deployment particularly important. Existing performance prediction methods often fail to address this type of shift effectively. We propose a novel two-stage performance prediction method specifically designed to tackle subpopulation shifts. I led the entire project.

**Prof. Peng Cui's lab**

*Research intern*

Tsinghua University

3 2023 — 7 2024

- **Domain Generalization**

From a data-centric perspective, we addressed a practical requirement of acquiring training samples from various domains on a limited budget to facilitate model generalization to target domain with distribution shift. To handle this, we introduced a Domain-wise Active Acquiring framework which iteratively optimizes the data acquisition strategy. I contributed some ideas and conducted all the experiments in this work.

- **Stable Learning**

Enhanced the sample-reweighting approach in stable learning by incorporating a loss function to emphasize the training of challenging samples. I took charge of curating datasets and plotting figures.

## PUBLICATIONS

---

- Yue He, **Dongbai Li** (co-first authors), Pengfei Tian, Han Yu, Jiashuo Liu, Hao Zou, Peng Cui. Domain-wise Data Acquisition to Improve Performance under Distribution Shift. **ICML 2024**.
- Han Yu, Yue He, Renzhe Xu, **Dongbai Li**, Jiayin Zhang, Wenchao Zou, Peng Cui. Sample Weight Averaging for Stable Prediction. **Under review**.
- **Dongbai Li**, Huan Zhang. SATE: A Two-Stage Approach for Performance Prediction in Subpopulation Shift Scenarios. **Under review**.

## SKILLS

---

- **Programming:** C++, Python
- **Machine learning:** principles, model utilization, data processing and analysis
- **Deep learning:** Pytorch

## SELECTED COURSES

---

- Fundamentals of Programming(C++) A
- Programming and Training(Python) A
- Probability and Statistics A

## REFERENCES

---

### **Prof. Peng Cui**

*Associate Professor(Tenured), Department of Computer Science and Technology, Tsinghua University*

E-mail: [cuip@tsinghua.edu.cn](mailto:cuip@tsinghua.edu.cn)

Scholar Profiles: [Personal Page](#) — [Google Scholar](#)

### **Prof. Huan Zhang**

*Assistant Professor, Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign*

E-mail: [huan@huan-zhang.com](mailto:huan@huan-zhang.com)

Scholar Profiles: [Personal Page](#) — [Google Scholar](#)