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 — 9 2022 — Present
B.E. in Computer Science and Technology — Cumulative GPA: 3.85

RESEARCH EXPERIENCE

Prof. Huan Zhang's lab — University of Illinois Urbana-Champaign (UIUC)
Visiting research intern — 7 2024 — Present

- **Benchmark Data Contamination(BDC)**
  We designed a systematic and controlled pipeline along with two novel metrics—fidelity and contamination resistance—to provide a fine-grained and comprehensive **assessment of existing BDC mitigation strategies**. Extensive experiments reveal that no existing strategy significantly improves resistance over the vanilla case (i.e., no benchmark update) across all benchmarks. I made significant contributions to conducting experiments and writing the article.

- **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 — Tsinghua University
Research intern — 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.

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**.

- Yifan Sun, Han Wang, **Dongbai Li** (co-first authors), Gang Wang, Huan Zhang. The Emperor's New Clothes in Benchmarking? A Rigorous Examination of Mitigation Strategies for LLM Benchmark Data Contamination. **Under review**.

SELECTED COURSES

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

REFERENCES

Prof. Peng Cui
Associate Professor(Tenured), Department of Computer Science and Technology, Tsinghua University
E-mail: 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
Scholar Profiles: Personal Page — Google Scholar