

JIANCHENG DONG

✉ jiancdong2-c@my.cityu.edu.hk · 🌐 dongjiancheng77 · 🏠 Homepage

🎓 EDUCATION

Nanjing University, Nanjing, China 2021 – 2025

B.S. in Artificial Intelligence

- **GPA:** 4.3/5.0 (86/100)

City University of Hong Kong, Hong Kong, China 2025 – Present

Ph.D. in Data Science

📄 PUBLICATIONS

[1] **Jiancheng Dong**, Lei Jiang, Wei Jin, Lu Cheng, *Threshold Filtering Packing for Supervised Fine-Tuning: Training Related Samples within Packs*, Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL Main Conference**) 2025. [paper]

[2] **Jiancheng Dong**, Pengyue Jia, Jingyu Peng, Maolin Wang, Yuhao Wang, Lixin Su, Xin Sun, Shuaiqiang Wang, Dawei Yin, Xiangyu Zhao, *Behavior-Equivalent Token: Single-Token Replacement for Long Prompts in LLMs*, in submission, ACL 2026 [paper]

[3] **Jiancheng Dong**, Pengyue Jia, Derong Xu, Jiawei Cheng, Jingyu Peng, Chao Zhang, Bowen Liu, Xin Sun, Lixin Su, Shuaiqiang Wang, Dawei Yin, Xiangyu Zhao, *How to Utilize Complementary Vision-Text Information for 2D Structure Understanding*, in submission, NeurIPS 2026

👥 EXPERIENCE

Baidu Inc., Search Science 5/2025 – 3/2026

Research Intern Advisors: Lixin Su, Shuaiqiang Wang, and Dawei Yin

- Focused on **context compression**; developed Behavior-Equivalent Token to equivalently replace complex long prompts, retaining 98% performance at compression ratios up to 3000×
- Expanded research to **visual-text compression**; leveraged the complementarity between textual and visual modalities to develop DiVA-Former, which improves upon the pure-text baseline by 23.9%.

University of Illinois Chicago, Responsible and Reliable AI Lab 3/2024 – 9/2024

Research Assistant Advisor: Lu Cheng and Wei Jin

- Developed the Threshold Filtering **Packing** method to reduce sequence cross-contamination, enhancing contextual learning and improving the utilization of training data.
- Conducted research in Trustworthy AI and LLM Fairness, investigating data-centric optimization strategies to mitigate demographic bias amplification during LLM **Alignment**.

Nanjing University, Natural Language Processing Group 4/2023 – 1/2024

Research Intern Advisor: Xinyu Dai and Zhen Wu

Explored **data selection** strategies for instruction tuning, aiming to optimize sample relevance and diversity while conducting preliminary experiments to identify methods for improving data efficiency.

♡ HONORS AND AWARDS

The School of Artificial Intelligence Scholarship for the 2024 Academic Year (4/413) March 2025

Outstanding Student Leader of Nanjing University May 2022

First Prize in E Fund Management Co., Ltd. 'AI+' Competition Oct 2023

People's Scholarship Dec 2023

International College Students' Innovation Competition Finals Dec 2023

Patent: LLM Knowledge Service Retrieval and Management System Jan 2024