Keyu Duan

+86 156 5258 5015; +65 8830 5701 | k.duan@u.nus.edu | GitHub | Portfolio @ kduan.live

Research Interests

My current research interests lie in the LLM post-training, reasoning and generalist agentic AI.

Education

National University of Singapore (NUS) — Singapore, Singapore
Ph.D. of Computer Science — *Aug. 2022 – May. 2026 (Expected)*

Beihang University — Beijing, CHINA
B.E. of Software Engineering — *Sept.2016 – Jun.2020*

Open-source Projects

GEM. A GYM for generalist agentic LLM. **Keyu Duan***, Zichen Liu*, Anya Sims*, Changyu Chen*. https://github.com/axon-rl/gem

Publications and Preprints

Keyu Duan, Zichen Liu, Xin Mao, Tianyu Pang, Changyu Chen, Qiguang Chen, Michael Qizhe Shieh, Longxu Dou. "Efficient Process Reward Model Training via Active Learning" *The Conference on Language Modeling 2025.*

Keyu Duan*, Yiran Zhao*, Zhili Feng, Jinjie Ni, Tianyu Pang, Qian Liu, Tianle Cai, Longxu Dou, Kenji Kawaguchi, Anirudh Goyal, Zico Kolter, Michael Qizhe Shieh. "Unnatural Languages Are Not Bugs but Features for LLMs" *The International Conference on Machine Learning 2025.*

Keyu Duan, Qian Liu, Tat-Seng Chua, Shuicheng Yan, Wei Tsang Ooi, Michael Qizhe Shieh, and Junxian He. "SimTeG: A frustratingly simple approach improves textual graph learning" *arXiv preprint arXiv:2308.02565 (2023).*

Qinggang Zhang*, **Keyu Duan***, Junnan Dong, Pai Zheng, Xiao Huang. "Logical Reasoning with Relation Network for Inductive Knowledge Graph Completion" *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.*

Junnan Dong, Qinggang Zhang, Xiao Huang, **Keyu Duan**, Qiaoyu Tan, and Zhimeng Jiang. "Hierarchy-aware multi-hop question answering over knowledge graphs" *Proceedings of the ACM Web Conference 2023, pp. 2519-2527. 2023.*

Keyu Duan, Zirui Liu, Peihao Wang, Wenqing Zheng, Kaixiong Zhou, Tianlong Chen, Xia Hu, and Zhangyang Wang. "A comprehensive study on large-scale graph training: Benchmarking and rethinking" *Advances in Neural Information Processing Systems (NeurIPS) Tack on Datasets and Benchmarks 35 (2022): 5376-5389.*

Tianlong Chen, Kaixiong Zhou, **Keyu Duan**, Wenqing Zheng, Peihao Wang, Xia Hu, and Zhangyang Wang. "Bag of tricks for training deeper graph neural networks: A comprehensive benchmark study" *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 45, no. 3 (2022): 2769-2781.*

Qinggang Zhang, Junnan Dong, **Keyu Duan**, Xiao Huang, Yezi Liu, and Linchuan Xu. "Contrastive knowledge graph error detection" *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM), pp. 2590-2599. 2022.*

Fuzhen Zhuang, Zhiyuan Qi, **Keyu Duan**, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. "A comprehensive survey on transfer learning" *Proceedings of the IEEE 109, no. 1 (2020): 43-76.*

Research Experience

SEA AI Lab Singapore, Singapore

Research Internship *1. 2025 – Now*

* Work on LLM reasoning and agentic AI.
* *ActPRM* (COLM 2025): We address the high annotation costs associated with training Process Reward Models (PRMs) by proposing ActPRM, an uncertainty-aware active learning framework that selectively annotates the most informative reasoning steps. By leveraging an ensemble PRM to estimate uncertainty and strategically labeling only uncertain data, ActPRM significantly reduces annotation costs while maintaining competitive performance. Extensive experiments demonstrate that ActPRM achieves a new state-of-the-art (75.0% on ProcessBench) with merely at most 20% of the labeling budget required by prior methods.

SoC at National University of Singapore Singapore, Singapore

Ph.D. student supervised by Prof. Michael Qizhe Shieh. *Aug. 2022 – Now*

* Work on large language models and its applications.
* *Unnatural Language (ICML 2025)* : We demonstrate that unnatural languages - strings that appears incomprehensible to humans but maintain semantic meanings for LLMs - contain latent features usable by models.
* *SimTeG (Preprint)*: We propose a frustratingly simple but effective framework for textual graph representation learning. An ensembled version of our framework achieved new SOTA performance on `OGBN-Arxiv`, one of the most prestigious graph benchmark. We empirically find that better text embeddings could bridge the gap among GNN models.

DATA Lab at Rice University remote

Ph.D. student supervised by Prof. Xia (Ben) Hu *Sept. 2020 – May. 2022*

* Work on Network Analytics and Graph Nerual Networks (GNNs), particularly focusing on seeking practical algorithms to build deep or scalable GNNs. Note: admitted as Ph.D. student but ceased due to VISA issue.
* *Bag of Tricks for Deep GNNs (TPAMI)*: we present the first fair and reproducible benchmark dedicated to assessing the "tricks" of training deep GNNs. We demonstrate that an organic combo of initial connection, identity mapping, group and batch normalization attains the new state-of-the-art results for deep GNNs on large datasets.
* *Scalable GNN benchmarking and Rethinking (NeurIPS dataset and benchmark track)*: we present an extensive benchmark study on large scale graph training with systematic taxonomy and hytperparamter unification. We investigate several aspects of scalable GNNs regarding effectiveness and efficiency. Besides, we propose an additional method that mitigates the large CPU consumption issue in precomputing-based GNNs.

DEEP Lab at Hong Kong Polytechnic University Hong Kong SAR, CHINA

Research Assistant supervised by Prof. Xiao Huang *Mar. 2021 – Mar. 2022*

* Work on knowledge graph reasoning and Completion.
* *Hierarchy-aware multi-hop question answering (WWW)*: We propose a tree-like question answering framework utilizing knowledge graphs.
* *Contrastive knowledge graph error detection (CIKM)*: We propose a error detection framework on knowledge graphs via contrastive learning.
* *IRENE: Infomax Relation Networks for Inductive Knowledge Graph Completion*: Arguing the limitations of embedding-based paradigm for knowledge graph completion, we propose a new modeling for KGC and a corresponding GNN framework to address inductive KGC.

Key Laboratory of Intelligent Information Processing, CAS Beijing, CHINA

Undergraduate Research Assistant supervised by Prof. Fuzhen Zhuang *May. 2019 – May. 2020*

* *A comprehensive survey on transfer learning (IEEE Proceedings)*: We provide a comprehensive survey on transfer learning, including an exhaustive taxonomy, benchmarks and its applications. Besides, I buit a transfer learning toolkit for primers that is easy-to-use for quick experiments.

Engineering Experience

Matrix and Its Application Laboratory at Beihang University Beijing, CHINA

Undergraduate Internship supervised by Prof. Hongyi Li *Sept. 2018 – Sept. 2019*

* *Automated Process of Satellite Reliability Assessment:* I participated in building an automated software platform to assess the reliability of satellite system components parallely.