

Advancing Language Models through Instruction Tuning: Recent Progress and Challenges

Zhihan Zhang, Renze Lou, Fangkai Jiao, Wenpeng Yin
and Meng Jiang

The capability of following instructions is a key dimension for AI systems. Therefore, in NLP, instruction tuning – the process of training language models to follow natural language instructions – has become a fundamental component of the model development pipeline. This tutorial addresses three critical questions within the field: (1) What are the current focal points in instruction tuning research? (2) What are the best practices in training an instruction-following model? (3) What new challenges have emerged? To answer these questions, the tutorial presents a systematic overview of recent advances in instruction tuning. It covers different stages in model training: supervised fine-tuning, preference optimization, and reinforcement learning. It introduces scalable strategies for building high-quality instruction data, explores approaches for training autonomous AI agents that handle complex real-world tasks, and discusses common criteria for evaluating instruction-following models. The audience will gain a comprehensive understanding of cutting-edge trends in instruction tuning and insights into promising directions for future research.

Zhihan Zhang, Applied Scientist, Amazon

email: zzhihan@amazon.com

website: <https://ytyz1307zzh.github.io>

Zhihan Zhang is an Applied Scientist at Amazon. He works on building intelligent AI agents powered by large language models for shopping applications. Zhihan earned his Ph.D. in Computer Science and Engineering from the University of Notre Dame, where his research centered around training instruction-following language models. Prior to that, Zhihan received his B.S. from Peking University. Zhihan has published over 30 papers in top NLP/ML conferences and journals, including ACL, EMNLP, ICLR, and NAACL.

Renze Lou, Ph.D. student, Department of Computer Science and Engineering, Pennsylvania State University

email: renze.lou@psu.edu

website: <https://renzelou.github.io>

Renze Lou is a third-year Ph.D. student at Pennsylvania State University. His research focuses on empowering AI agents to assist in various professional domains. He has extensive research experience in instruction tuning and following, agentic systems, and AI4Research. Renze has (co-)authored papers at top-tier conferences, including ICLR, ICML, AACL, ACL, and EMNLP. He has also completed research internships at Salesforce Research and Microsoft Research.

Fangkai Jiao, Ph.D. student, College of Computing and Data Science, Nanyang Technological University

email: fangkai002@e.ntu.edu.sg

website: <https://jiaofangkai.com>

Fangkai Jiao is a fourth-year PhD student at Nanyang Technological University and the Institute of Infocomm Research, A*STAR, Singapore. Prior to his PhD, he received his M.Eng. and B.Eng. from Shandong University. Fangkai's research focuses on weak-supervised training and data synthesis for machine reasoning and large language models. He has published several papers in top-tier conferences and journals, including ACL, EMNLP, ICLR, and TPAMI. He has also held research internships at DAMO Academy, Alibaba Group, Microsoft Research Asia, and Bytedance Seed Team.

Wenpeng Yin, Assistant Professor, Department of Computer Science and Engineering, Penn State University

email: wenpeng@psu.edu

website: <https://www.wenpengyin.org>

Wenpeng Yin is an Assistant Professor in the Department of Computer Science and Engineering at Penn State University, USA. Dr. Yin is working on AI to automate NLP research. He has experience presenting tutorials at ACL 2023, EMNLP 2023, KONVENS 2023, and EMNLP 2024. He led the workshop "WISE-Supervision" co-located with AKBC 2022 and the 1st and 2nd AI4Research Workshops at IJCAI 2024/AAAI 2025.

Meng Jiang, Associate Professor and Frank M. Freimann Collegiate Professor of Computer Science and Engineering, University of Notre Dame

email: mjiang2@nd.edu

website: <http://www.meng-jiang.com>

Meng Jiang is an Associate Professor and Frank M. Freimann Collegiate Professor of Computer Science and Engineering at the University of Notre

Dame. He is appointed as the Director of Foundation Models at the Lucy Family Institute for Data and Society as well as the Program Chair of ND-IBM Tech Ethics Lab. He is an Amazon Scholar. His research interests are data mining, machine learning, and natural language processing (NLP) for applications such as material discovery, recommender system, question answering, education, and mental health. His recent projects focus on knowledge-augmented NLP, instructed large language model (LLM), self-correct LLM, personalized LLM, unlearned LLM, graph data augmentation, and graph diffusion model. He has delivered 15 conference tutorials and organized ten workshops on these topics. He has received multiple best paper awards and awarded NSF CAREER.