

JINJIE NI

Email: jinjieni@nus.edu.sg | Homepage: <https://jinjie.one>

Research Interests

Large (multi-modal) language models; Compute-optimal learning; System-algorithm co-design.

Experiences

Academia

- National University of Singapore** 2023 - present
Research Fellow
- Foundation Models.
- Nanyang Technological University** 2020 - 2023
Ph.D. in Computer Science
- Efficient Language Models and Dialogue Systems.
- Harvard University**, Institute for Applied Computational Science
Research Assistant (remote)
- VAE-GAN variants.
- Northwestern Polytechnical University** 2016 - 2020
B.Eng. in Electrical Engineering
- Multimodal Models.

Industry

- Research Intern at Alibaba Group, Singapore** April 2022 - Oct 2022
DAMO Academy
- In charge of modality alignment for pre-trained models. Worked with Dr. Yukun Ma.
- Research Intern at Continental** Sept 2020 - March 2022
Continental-NTU Corp Lab
- In charge of fusing task-oriented and open-domain dialogue agents. Worked with Dr. Rui Mao.
- Research Intern at Chinese Academy of Sciences** Feb 2020 - June 2020
Institute of Automation
- In charge of anchor-free position estimation and object detection. Worked with Dr. Sen Xin.
- Institute of Computing Technology Oct 2018 - Nov 2018
- Training abstractive summarization models. Worked with Dr. Shuai Jiao.

Featured Research

For full publication list, see [Google Scholar](#).

- **MixEval-X**
 - MixEval-X: Any-to-Any Evaluations from Real-World Data Mixtures. **ICLR 2025**. [[Twitter](#)]
 - **Jinjie Ni**, Yifan Song, Deepanway Ghosal, Bo Li, David Junhao Zhang, Xiang Yue, Fuzhao Xue, Zian Zheng, Kaichen Zhang, Mahir Shah, Kabir Jain, Yang You, Michael Qizhe Shieh.
 - MixEval-X is the **first** any-to-any, real-world benchmark featuring **diverse input-output modalities, real-world task distributions, consistent high standards across modalities, and dynamism**. It achieves up to **0.98** correlation with arena-like multi-modal evaluations while being way more efficient.
- **MixEval**
 - MixEval: Deriving Wisdom of the Crowd from LLM Benchmark Mixtures. **NeurIPS 2024** main track (poster). [[Twitter](#)]
 - **Jinjie Ni**, Fuzhao Xue, Xiang Yue, Yuntian Deng, Mahir Shah, Kabir Jain, Graham Neubig, Yang You.

- Building golden-standard LLM evaluation from off-the-shelf benchmark mixtures. The **best** LLM evaluation at the time of release for its **SOTA** model ranking accuracy (0.96 correlation with Chatbot Arena) and efficiency (6% the time and cost of running MMLU). Moreover, it's dynamic.
- **OpenMoE**
 - OpenMoE: An Early Effort on Open Mixture-of-Experts Language Models. **ICML 2024** (poster). [\[Twitter\]](#)
 - Fuzhao Xue, Zian Zheng, Yao Fu, **Jinjie Ni**, Zangwei Zheng, Wangchunshu Zhou, Yang You.
 - The **first fully open** MoE-based Decoder-only LLM trained over chinchilla scaling law.
- **InstructWild**
 - Instruction in the Wild: A User-Based Instruction Dataset. Github.
 - **Jinjie Ni**, Fuzhao Xue, Yuntian Deng, Jason Phang, Kabir Jain, Mahir Hitesh Shah, Zangwei Zheng, Yang You.
 - The **first** large-scale instruction tuning dataset harvested from the web.
- **GHA**
 - Finding the Pillars of Strength for Multi-Head Attention. **ACL 2023** main track (poster).
 - **Jinjie Ni**, Rui Mao, Zonglin Yang, Han Lei, Erik Cambria.
 - Cutting off redundancy for Transformer layers. **SOTA** efficiency and performance among efficient transformers. Concurrent work of GQA, cited and discussed in the GQA paper.
- **PAD**
 - Adaptive Knowledge Distillation between Text and Speech Pre-trained Models. **ICASSP 2023** (oral).
 - **Jinjie Ni**, Yukun Ma, Wen Wang, Qian Chen, Dianwen Ng, Han Lei, Trung Hieu Nguyen, Chong Zhang, Bin Ma, Erik Cambria.
 - Knowledge distillation between text and speech pre-trained models. The **SOTA** text-speech distillation method at the time of release.
- **HiTKG**
 - HiTKG: Towards Goal-Oriented Conversations via Multi-Hierarchy Learning. **AAAI 2022** (oral).
 - **Jinjie Ni**, Vlad Pandealea, Tom Young, Haicang Zhou, Erik Cambria.
 - The **first** work that trains agents to actively guide the conversations. It ushers in a **new era** of intelligence for dialogue agents. The **SOTA** approach for turn-level dialogue reasoning tasks.
- **FusedChat**
 - FusedChat: Towards Fusing Task-Oriented Dialogues and Chitchat in Multi-turn Conversational Agents. **AAAI 2022** (oral).
 - Tom Young, Frank Xing, Vlad Pandealea, **Jinjie Ni**, Erik Cambria.
 - The **first** attempt of fusing task-oriented and open-domain dialogue systems.
- **Recent Advances in Deep Learning Based Dialogue Systems**
 - Recent Advances in Deep Learning Based Dialogue Systems. **AIRE**.
 - **Jinjie Ni**, Tom Young, Vlad Pandealea, Fuzhao Xue, Erik Cambria.
 - An 80-page systematic review for dialogue systems. One of the **most** cited dialogue system reviews.

Services

Conference PC Member / Reviewer

- ICML 2025, ICLR 2025, Neurips 2024, ACL 2024, EMNLP 2024, ACL 2023, EMNLP 2023, AAAI 2023

Journal Reviewer

- Knowledge-Based Systems, Information Fusion, Artificial Intelligence Review, Cognitive Computation

Co-organizer

- MLNLP community