Kailai Yang, B.Eng.

✉ kailai.yang@manchester.ac.uk   𝔾 Google Scholar   in Linkedin
🌐 Homepage   ⌨ Github
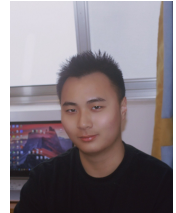
Education

2017 – 2021	**B.Eng. Computer Science, Harbin Institute of Technology**, China Thesis title: *Knowledge-Interactive Network with Sentiment Polarity Intensity-Aware Multi-Task Learning for Emotion Recognition in Conversations* GPA: *88/100* Class ranking: *1/32*
2022 – 2026	**Ph.D. Computer Science, The University of Manchester**, UK Advisors: *Prof. Sophia Ananiadou* and *Dr. Junichi Tsujii*. Funding Source: *President's Doctoral Scholar Award* Research Interests: *Language Model Alignment, Efficient Fine-tuning of LLMs, Computational Social Science*

Employment History

Nov, 2022 – Aug, 2023	**Research Intern**, Artificial Intelligence Research Centre, National Institute of Advanced Industrial Science and Technology (AIST), Japan.
Jun, 2020 – Nov, 2020	**Research Intern**, UC Irvine, School of Information and Computer Sciences, USA.
Dec, 2020 – July, 2021	**Research Intern**, Harbin Institute of Technology, China.
Feb, 2022 – May, 2024	**Teaching Assistant**, The University of Manchester, UK.

Selected Publications

For a full list of my publications, please visit my Google Scholar page.

Conference Proceedings

1. K. Yang, Z. Liu, Q. Xie, J. Huang, T. Zhang, and S. Ananiadou, "Metaaligner: Towards generalizable multi-objective alignment of language models," in *arXiv preprint arXiv:2403.17141*, 2024.

2. K. Yang, T. Zhang, Z. Kuang, Q. Xie, J. Huang, and S. Ananiadou, "Mentallama: Interpretable mental health analysis on social media with large language models," in *Proceedings of the ACM on Web Conference 2024*, ser. WWW '24, , Singapore, Singapore, Association for Computing Machinery, 2024, pp. 4489–4500, ISBN: 9798400701719. 🔗 DOI: 10.1145/3589334.3648137.

3. K. Yang, S. Ji, T. Zhang, Q. Xie, Z. Kuang, and S. Ananiadou, "Towards interpretable mental health analysis with large language models," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, H. Bouamor, J. Pino, and K. Bali, Eds., Singapore: Association for Computational Linguistics, Dec. 2023, pp. 6056–6077. 🔗 DOI: 10.18653/v1/2023.emnlp-main.370.

4. K. Yang, T. Zhang, S. Ji, and S. Ananiadou, "A bipartite graph is all we need for enhancing emotional reasoning with commonsense knowledge," in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, ser. CIKM '23, , Birmingham, United Kingdom, Association for Computing Machinery, 2023, pp. 2917–2927, ISBN: 9798400701245. 🔗 DOI: 10.1145/3583780.3614758.

5 K. Yang*, Y. Xie*, C. Sun, B. Liu, and Z. Ji, "Knowledge-interactive network with sentiment polarity intensity-aware multi-task learning for emotion recognition in conversations," in *Findings of the Association for Computational Linguistics: EMNLP 2021*, M.-F. Moens, X. Huang, L. Specia, and S. W.-t. Yih, Eds., Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 2879–2889. 🔗 DOI: 10.18653/v1/2021.findings-emnlp.245.

Journal Articles

1 K. Yang, T. Zhang, H. Alhuzali, and S. Ananiadou, "Cluster-level contrastive learning for emotion recognition in conversations," *IEEE Transactions on Affective Computing*, 2023.

2 K. Yang, T. Zhang, and S. Ananiadou, "Disentangled variational autoencoder for emotion recognition in conversations," *IEEE Transactions on Affective Computing*, 2023.

3 K. Yang, T. Zhang, and S. Ananiadou, "A mental state knowledge–aware and contrastive network for early stress and depression detection on social media," *Information Processing & Management*, vol. 59, no. 4, p. 102 961, 2022.

Skills

Languages	Strong reading, writing, and speaking competencies in English and Mandarin Chinese.
Coding	Python, LaTeX, C, Git.
OS	Linux, Windows, MacOS.
Library	Pytorch, Transformers, Deepspeed, .
LLMs	Continual pre-training, Supervised Fine-tuning, RLHF (PPO), DPO, In-context Learning, etc.

Achievements

Open-sourced Projects

2023   **MentaLLaMA**: the first interpretable mental health analysis large language model series; Github stars: 160.

PIXIU: the first financial large language models (LLMs); Github stars: 414.

Academic Service

Conference Program Committee/Reviewer: NAACL' 24, WWW' 24, ACL' 23, 24, EACL' 24, NLPCC' 24, BIBM' 22, PAKDD' 22, 23

Journal Reviewer: IEEE Transactions on Knowledge and Data Engineering, ACM Transactions on Computing for Healthcare, CAAI Transactions on Intelligence Technology, Artificial Intelligence Review, Information Processing & Management, Scientific Reports.

Organizer

FinNLP-AgentScen@IJCAI-2024.

CIKM' 23 Session Chair.

Ellis PhD Program Evaluator, 2023.

The **Multimodal Large Language Model Talk**.