

Education	<b>The University of Texas at Dallas</b>	2024 - present
	- Ph.D. in Computer Science	Dallas, TX, USA
	- Advisor: <a href="#">Zhiyu Zoey Chen</a>	
	<b>Virginia Tech</b> (transferred out)	2023 - 2024
	- Ph.D. in Computer Science	Blacksburg, VA, USA
	<b>Soochow University</b>	2020 - 2023
	- M.S. in Computer Science	Suzhou, Jiangsu, China
	- Advisors: <a href="#">Wenliang Chen</a> , <a href="#">Xiabing Zhou</a>	
	- Outstanding Graduate; Outstanding Thesis	
	<b>Nanjing University of Posts and Telecommunications</b>	2016 - 2020
	- B.Eng. in Computer Science	Nanjing, Jiangsu, China
	- GPA: 3.96 / 5 (Top 3% in the Department)	

Research Summary	My current research focuses on enhancing (large) language models for social good, with a particular emphasis on <i>health domains</i> [1, 2, 4]. I am actively working on <b>post-training techniques</b> ( <i>reinforcement learning from human feedback</i> ) [1, 6], and <b>synthetic data generation</b> [1, 7] to equip models with complex and reliable capabilities, such as <i>instruction following</i> and <i>personalization</i> . I have experience in evaluating and improving language models, particularly in their <b>dialogue abilities</b> , such as keeping <i>self-consistent</i> [5] and <i>safe</i> [6] in conversation or understanding speaker's <i>core semantic meaning</i> [7] and <i>emotion</i> [8]. Additionally, I have a strong interest in vision-language models, retrieval-augmented methods, and language agents. Totally, I have 9 papers along these directions including 7 first-authored papers and 7 peer-reviewed publications in top-tier AI conferences such as ACL, NAACL, NeurIPS, etc.	
------------------	---	--

Research Experiences	<b>Human Language Technology Research Institute at UTD</b>	Sep. 2024 - present
	<i>Student Researcher</i> (Advisor: <a href="#">Zhiyu Zoey Chen</a> )	Dallas, TX
	Topics: Preference Learning, Personalization	
	<b>Tencent AI Lab</b>	May. 2024 - Jul. 2024
	<i>Research Intern</i> (Mentors: <a href="#">Haitao Mi</a> , <a href="#">Linfeng Song</a> )	Seattle, WA
	Topics: Reinforcement Learning from Human Feedback	
	<b>Tencent AI Lab</b>	Dec. 2021 - Dec. 2022
	<i>Research Intern</i> (Mentor: <a href="#">Linfeng Jin</a> )	Shenzhen, China
	Topics: Dialogue Understanding, Semi-supervised Learning	
	<b>Institute of Human Language Technology at Soochow University</b>	Sep. 2020 - Jul. 2021
	<i>Student Researcher</i> (Advisors: <a href="#">Wenliang Chen</a> , <a href="#">Xiabing Zhou</a> )	Suzhou, China
	Topics: Dialogue State Tracking, Conversational Emotion Recognition	

Papers	(*denotes equal contribution)	
	1. Aligning Large Language Models to Assist in Psychotherapy ( <i>Ongoing</i> )	
	<b>Mian Zhang</b> , Shaun M. Eack, Zhiyu Chen.	
	2. <a href="#">CBT-Bench: Evaluating Large Language Models on Assisting Cognitive Behavior Therapy</a>	
	(NAACL 2025)	

**Mian Zhang\***, Xianjun Yang\*, Xinlu Zhang, Travis Labrum, Jamie C. Chiu, Shaun M. Eack, Fei Fang, William Yang Wang, Zhiyu Chen.

3. [IDEA: Enhancing the Rule Learning Ability of Large Language Model Agent through Induction, Deduction, and Abduction](#) (*NeurIPS 2024 Workshop on Open-World Agents, ACL under submission*)  
Kaiyu He, **Mian Zhang**, Shuo Yan, Peilin Wu, Zhiyu Zoey Chen.
4. [Large Language Models for Disease Diagnosis: A Scoping Review](#) (*Preprint*)  
Shuang Zhou\*, Zidu Xu\*, **Mian Zhang\***, Chunpu Xu\*, Yawen Guo, Zaifu Zhan, Sirui Ding, Jiashuo Wang, Kaishuai Xu, Yi Fang, Liqiao Xia, Jeremy Yeung, Daochen Zha, Mingquan Lin, Rui Zhang.
5. [Inconsistent dialogue responses and how to recover from them](#) (*EACL 2024*)  
**Mian Zhang**, Lifeng Jin, Linfeng Song, Haitao Mi, Dong Yu.
6. [SafeConv: Explaining and Correcting Conversational Unsafe Behavior](#) (*ACL 2023 Oral*)  
**Mian Zhang**, Lifeng Jin, Linfeng Song, Haitao Mi, Wenliang Chen, Dong Yu.
7. [Friend-training: Learning from Different but Related Tasks](#) (*EACL 2023*)  
**Mian Zhang**, Lifeng Jin, Linfeng Song, Haitao Mi, Xiabing Zhou, Dong Yu.
8. [Emotion Recognition in Conversation from Variable-Length Context](#) (*ICASSP 2023*)  
**Mian Zhang**, Xiabing Zhou, Wenliang Chen, Min Zhang.
9. [A Pairing Enhancement Approach for Aspect Sentiment Triplet Extraction](#) (*KSEM 2023*)  
Fang Yang, **Mian Zhang**, Gongzhen Hu, Xiabing Zhou.

Skills

*Programming Language:* Python, C/C++, Shell,  $\text{\LaTeX}$ , Matlab

*Machine Learning Framework:* PyTorch, NumPy, Transformers, Scikit-learn

*Tool & Software:* Vim, Git, pdb, Matplotlib, Pandas

*Natural Language:* Mandarin (native), English (advanced)

Selected Awards

Outstanding Master's Thesis	2024
Outstanding Graduate Award at Soochow University	2023
First-class Scholarship at Soochow University	2020 - 2021
Principal's Award	2019
<a href="#">CUMCM National First Prize (top 1%)</a>	2018
First-class Scholarship at NJUPT	2017 - 2019