(last update: 2024.5)

Haolin Deng

➡ hldeng028@gmail.com

A Homepage

Github Google Scholar

RESEARCH INTERESTS

Natural Language Processing; Machine Learning

I am currently working on grounding and evaluating large language models in complex real-world applications. I have also worked in the following areas: retrieval-augmented generation; text summarization; information extraction; dialog generation; commonsense reasoning.

EDUCATION

Harbin Institute of Technology, Shenzhen, China	2022.9-present
Master of Science in Computer Science and Technology, supervised by Prof. Ruifeng Xu.	
Harbin Institute of Technology, Shenzhen, China	2018.9-2022.6
Bachelor of Engineering in Computer Science and Technology	
University of California, Berkeley, US	2019.6-2019.8
Visiting student	

EXPERIENCE

WeChat Web search team, Tencent

2024.3-present

Research Intern

• Exploring the potential of AI search engine driven by large language models.

QB research team, Tencent

2023.5-2024.3

Research Intern

- Explored the potential of AI search engine driven by large language models.
- One long paper on attributed language models accepted to ACL 2024 main conference.
- Participated in improving the detection of time-sensitive user queries in QQ Browser.

QB research team, Tencent

2022.3-2022.6

Research Intern

- Explored algorithms for automatically extracting event threads for trending news.
- One long paper on open event extraction accepted to EMNLP 2022 main conference.

PUBLICATION

In Conference Proceedings

- Haolin Deng, Chang Wang, Xin Li, Dezhang Yuan, Junlang Zhan, Tianhua Zhou, Jin Ma, Jun Gao, Ruifeng Xu. WebCiteS: Attributed Query-Focused Summarization on Chinese Web Search Results with Citations. Accepted to the 62nd Annual Meeting of the Association for Computational Linguistics (ACL), 2024.
- Haolin Deng, Yanan Zhang, Yangfan Zhang, Wangyang Ying, Changlong Yu, Jun Gao, Wei Wang, Xiaoling Bai, Nan Yang, Jin Ma, Xiang Chen, and Tianhua Zhou. Title2Event: Benchmarking Open Event Extraction with a Large-scale Chinese Title Dataset. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2022.
- Jun Gao, Yuhan Liu, Haolin Deng, Wei Wang, Yu Cao, Jiachen Du, and Ruifeng Xu. Improving Empathetic Response Generation by Recognizing Emotion Cause in Conversations. In Findings of the Association for Computational Linguistics: EMNLP, 2021.

SKILLS

Programming Python, PyTorch, C, C++

English TOEFL 106/120

AWARDS

First prize in Annual University Academic Scholarship, HITSZ	2022-2023
2nd place in Shenzhen Invitational Debate Tournament for College Students	2022
Special Prize in 2020 National English Competition for College Students	2020
ACADEMIC SERVICE	
Reviewer ACL-ARR (2023, 2024)	
TEACHING	

TEACHING

Teaching assistant for undergraduate course Computer Program Design, HITSZ	2023 spring
Teaching assistant for undergraduate course Foundations of Data Science, HITSZ	2021 spring