Huanhuan Ma

CONTACT INFORMATION

huanhuanma.top | Google Scholar
andy.huanhuan.ma@gmail.com | GitHub
Beijing, China | +86 13393906361

EDUCATION

University of Chinese Academy of Sciences, Beijing, China — Sep 2021 – July 2024
Master in Artificial Intelligence, GPA: 3.51/4.00

- Advisors: Prof. Liang Wang, Prof. Qiang Liu
- Thesis: Explainable Automated Fact Verification Research
- Award: Ministry of Education - Huawei "Intelligent Base" Future Star — 2022

Zhengzhou University, Zhengzhou, China — Sep 2016 – July 2020
Bachelor in Software Engineering

RESEARCH INTERESTS

My research primarily focuses on advancing Responsible AI, with an emphasis on:

- **Large model evaluation:** a) Traditional task-based evaluation [6][3], b) Psychological-based behavior evaluation of models [1]
- **Large model capability enhancement:** Targeted strengthening of model weaknesses through synthetic data generation

RESEARCH EXPERIENCE

Evaluation of Psychological Traits in LLMs — December 2023 – Present
NC State University, GIC Lab, NC, USA
Intern Researcher

- Leading a project to develop a more robust psychological behavior evaluation tool for assessing large language models (LLMs), with a manuscript submitted to ICLR 2025 [1]

Synthetic Data Generation — July 2024 – Present
Beijing Academy of Artificial Intelligence, (BAAI), Beijing, China
Research Intern

- Contributed to generating high-quality synthetic data for model training and identifying model weaknesses

Fact-Checking — August 2022 - March 2024
Institute of Automation, Chinese Academy of Sciences, CRIPAC, Beijing, China
Master Student

- Proposed a novel explainable fact-checking dataset with textual explanations [3]
- Developed an innovative method for explainable multi-modal misinformation detection [4]
- Introduced a framework to address the weaknesses of large language models in fact-checking tasks [2]

LLM Knowledge Editing — July 2023 - March 2024
Institute of Automation, Chinese Academy of Sciences, CRIPAC, Beijing, China
Master Student

- Investigated entity and relational knowledge differences through knowledge localization [5]

PRE-PRINTS

[1] **Huanhuan Ma**, Haisong Gong, Xiaoyuan Yi, Xing Xie, Dongkuan Xu. "Leveraging Implicit Sentiments: Enhancing Reliability and Validity in Psychological Trait Evaluation of LLMs" **(Under-review)**. [PDF],[OpenReview]

[2] Haisong Gong, **Huanhuan Ma**, Qiang Liu, Shu Wu, Liang Wang. "Navigating the Noisy Crowd: Finding Key Information for Claim Verification." **(Arxiv, July 2024)**. [PDF]

Peer-reviewed Publications

[3] **Huanhuan Ma**, Weizhi Xu, Yifan Wei, Liuji Chen, Liang Wang, Qiang Liu, Shu Wu, Liang Wang. "EX-FEVER: A Dataset for Multi-hop Explainable Fact Verification." *Findings of the Association for Computational Linguistics* **(ACL 2024 Findings)**. [PDF], [Code], [Poster], [Slides]

[4] **Huanhuan Ma**[*], Jinghao Zhang[*], Qiang Liu, Shu Wu, Liang Wang. "Interpretable Multimodal Out-of-Context Detection with Soft Logic Regularization." *IEEE International Conference on Acoustics, Speech and Signal Processing* **(ICASSP 2024 Oral)**. [PDF], [Slides]

[5] Yifan Wei, Xiaoyan Yu, Yixuan Weng, **Huanhuan Ma**, Yuanzhe Zhang, Jun Zhao, Kang Liu. "Does Knowledge Localization Hold True? Surprising Differences Between Entity and Relation Perspectives in Language Models." *ACM International Conference on Information and Knowledge Management* **(CIKM 2024 Short)**. [PDF], [Code]

[6] Yifan Wei, Yisong Su, **Huanhuan Ma**, Xiaoyan Yu, Fangyu Lei, Yuanzhe Zhang, Jun Zhao, Kang Liu. "MenatQA: A New Dataset for Testing the Temporal Comprehension and Reasoning Abilities of Large Language Models." *Findings of the Association for Computational Linguistics* **(EMNLP 2023 Findings)**. [PDF], [Code]

[7] Liping Wang, Qiang Liu, **Huanhuan Ma**, Shu Wu, Liang Wang. "Multi-Cause Learning for Diagnosis Prediction." *International Conference on Data Mining and Big Data* **(DMBD 2022)**. [PDF]

Projects

Awesome-LLM-based-Evaluators [GitHub] ★28
- Curated a comprehensive list of LLM-based evaluators for various NLP tasks

Internships

NC State University, GIC Lab, NC, USA December 2023 – Present
Intern Researcher
- Leading a project to develop a more robust psychological behavior evaluation tool for assessing large language models (LLMs), with a manuscript submitted to ICLR 2025 [1]

Beijing Academy of Artificial Intelligence, BAAI, Beijing, China July 2024 – Present
Research Intern
- **Synthetic Data Generation**: Contributed to generating high-quality synthetic data for model training and identifying model weaknesses
- **Infinity-MM Project**: Contributed to the optimization of training data for enhancing large vision-language model

Academic Services

Program Committee
- The International Conference on Learning Representations (ICLR) 2025
- ACM International Conference on Information and Knowledge Management (CIKM) 2024
- Annual Conference on Neural Information Processing Systems (NeurIPS) Dataset & Benchmark track 2023

Skills

Language: Chinese (Native), English (Fluent)
Programming: Python, LaTeX, SQL, HTML/CSS
Technologies: Git, Linux
Libraries: PyTorch, VLLM, LiteLLM, Faiss, SQLite