# Yuzhe (Toby) Yang

yuzheyang@link.cuhk.edu.cn yuzheyang.com Google Scholar

Last updated: September, 2025

## Research Interest

My research focuses on building reliable and trustworthy AI systems that bridge the gap between machines and the real world. I am particularly interested in **Human-AI Interaction** [C2], **Trustworthy NLP** [C3] and **Socially Aware NLP** [C1]. My vision is to enhance the social intelligence of (vision) language models, empowering them to not only understand the principles of the physical world but also to gain insight into complex social environments, enabling reliable and meaningful interactions with humans.

## Education

**The Chinese University of Hong Kong, Shenzhen** Sep. 2021 – May 2025
B.Eng. in Computer Science & Engineering Shenzhen, China

## Publications

(∗ indicates equal contribution)

### Conference & Workshop Papers

[C1] **Yuzhe Yang***, Yifei Zhang*, Minghao Wu*, Kaidi Zhang, Yunmiao Zhang, et al. *TwinMarket: A Scalable Behavioral and Social Simulation for Financial Markets*. 2025. URL: https://arxiv.org/abs/2502.01506.
**Best Paper Award, Travel Grant Award** *(ICLR 2025 Workshop on Advances in Financial AI)*
Under Review at *NeurIPS 2025*.

[C2] **Yuzhe Yang***, Yifei Zhang*, Yan Hu*, Yilin Guo, Ruoli Gan, et al. "UCFE: A User-Centric Financial Expertise Benchmark for Large Language Models". In: *Findings of the Association for Computational Linguistics: NAACL 2025 (Findings of NAACL)*. 2025, pp. 5429–5448. URL: https://aclanthology.org/2025.findings-naacl.300/.

[C3] Zihao Li, Xu Wang, **Yuzhe Yang**, Ziyu Yao, Haoyi Xiong, et al. "Feature Extraction and Steering for Enhanced Chain-of-Thought Reasoning in Language Models". In: *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2025. URL: https://arxiv.org/abs/2505.15634.

[C4] Jiaqi Wu, Simin Chen, Jing Tang, **Yuzhe Yang**, Yiming Chen, et al. "FDPT: Federated Discrete Prompt Tuning for Black-Box Visual-Language Models". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 2025.

### Journal Papers

[J1] Chi Li, Xixian Qi, **Yuzhe Yang**, Zhuo Zeng, Lianmin Zhang, et al. "FAST-CA: Fusion-based Adaptive Spatial–Temporal Learning with Coupled Attention for airport network delay propagation prediction". In: *Information Fusion* 107.1 (2024), p. 102326. URL: https://www.sciencedirect.com/science/article/pii/S1566253524001040.

### Preprints & Technical Reports

[P1] Jimin Huang, Mengxi Xiao, Dong Li, Zihao Jiang, **Yuzhe Yang**, et al. *Open-FinLLMs: Open Multimodal Large Language Models for Financial Applications*. 2025. URL: https://arxiv.org/abs/2408.11878.

[P2] Jiaqi Wu, Simin Chen, **Yuzhe Yang**, Yijiang Li, Shiyue Hou, et al. *FedDTPT: Federated Discrete and Transferable Prompt Tuning for Black-Box Large Language Models*. 2024. URL: https://arxiv.org/abs/2411.00985.

## Experiences

**UCSB ERIC Lab, UCSB NLP Group** Jul. 2025 – Present
Visiting Student (Advisor: Prof. Xin Eric Wang) Santa Barbara, CA, USA
Topic: Agent, Trustworthy NLP

**CUHK-Shenzhen NLP Group** Jun. 2024 – Present
Undergraduate Research Assistant (Advisors: Prof. Benyou Wang, Dr. Yan Hu) Shenzhen, China

• **Multi-Agent System** [C1, *ICLR W 25'*]: Developed a scalable multi-agent framework to simulate financial market. It successfully replicated real-world social emergent phenomena, such as financial bubbles and volatility clustering, by modeling the interaction of hundreds of LLM-powered investors.
• **Human-AI Interaction** [C2, *NAACL 25'*]: Evaluating LLMs on complex, real-world financial tasks by designing a user-centric framework with dynamic, multi-turn interactions.
• **Trustworth NLP** [C3, *EMNLP 25'*]: Pioneered a feature steering framework to extract pure reasoning features from LLMs by separating verbal and symbolic processes, and developed a "SAE-free" algorithm to enhance mathematical reasoning capabilities without external data and training.

**TheFinAI** Jun. 2024 – Oct. 2024
Researcher Remote

• **Financial (Vision) Language Model** [P1]: Enhanced a Llama-3-8B model for financial applications by conducting continued pre-training (Fin-Llama) on a massive financial corpus and developing its multi-modal extension (Fin-Llava) for advanced tabular understanding and chart reasoning.

**School of Data Science, CUHK-Shenzhen** Aug. 2023 – Jun. 2024
Undergraduate Research Assistant (Advisor: Prof. Jianfeng Mao) Shenzhen, China

• **Spatial-Temporal Modeling** [J1, *Information Fussion 24'*]: Designed and implemented a GNN model, which integrates adaptive graph learning, coupled attention, and periodicity feature extraction to provide a comprehensive analysis of the interplay between departure/arrival delays and spatial-temporal correlations in airport networks.

## Awards

**Best Paper Award** (ICLR 2025 Workshop on Advances in Financial AI) 2025
**Travel Grant Award** (ICLR 2025 Workshop on Advances in Financial AI) 2025
**Kaggle Silver Medal** (AI Mathematical Olympiad - Progress Prize 2) 2025
**Undergraduate Research Award** (CUHK-Shenzhen) 2024, 2025
**Outstanding College Contribution Award** (CUHK-Shenzhen) 2021, 2022

## Presentations

**TwinMarket: A Scalable Behavioral and Social Simulation for Financial Markets**

• Guest lecture for CSC6052, Spring 2025, CUHK-Shenzhen Apr. 2025
• Contributed talk at the ICLR 2025 Workshop, Singapore Apr. 2025
• Invited talk at the Wisemodel Open-source Series (Virtual) May 2025

## Projects

**Quant-GPT: Money is All You Need** | PyTorch, Transformers, ChromaDB May 2024
*Multi-Agent system for A-share market investment decisions.*

• Fine-tuned LLM with sentiment analysis and real-world market data integration
• Implemented RAG and multi-agent systems for dynamic financial news synthesis
• Achieved Sharpe Ratio: 0.40, Annualized Return: 7.26%, Max Drawdown: 13.61%

**Travel Insurance Recommendation AI System** | PyTorch, LangChain Apr. 2024

*Spatial-Temporal model for flight delay prediction and personalized travel insurance recommendation.*
• Fine-tuned LLM using insurance corpus for domain-specific question-answering
• Achieved 83% accuracy in identifying user intent for insurance recommendations
• Integrated GNN-based deep learning and LLM agents for delay prediction and sentiment assessment

## Skills

**Programming Languages:** Python, C/C++, HTML/CSS
**Developer Tools:** Pytorch, Transformers, LangChain, Faiss, Git, Slurm

## Services

**Reviewer:** IJCAI 2025, ICLR 2025 Workshop, ACL 2025 SRW
**Organizer:** Reading Seminar at CUHK-Shenzhen

References available upon request.