Yuzhe (Toby) Yang

Research Assistant, 2001 Longxiang Avenue,
Longgang District, Shenzhen, 518172, China

(+86) 183-1076-2536
yuzheyang@link.cuhk.edu.cn
tobyyang7.github.io
Google Scholar

Education

University of California, Santa Barbara — Santa Barbara, CA, USA
Visiting Student — *Jun. 2025 – Oct. 2025*

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

Research Interests

Language Model: Human-Machine Interaction, Agent System, AI for Scientific Applications, Trustworthy NLP
Data Mining: Social Computing, Spatial-Temporal Modeling

Publications

TwinMarket: A Scalable Behavioral and Social Simulation for Financial Markets [online]
Yuzhe Yang*, Yifei Zhang*, Minghao Wu*, Kaidi Zhang, Yunmiao Zhang, Honghai Yu, Yan Hu, Benyou Wang
ICLR 2025 Workshop on Advances in Financial AI, Best Paper Award.
Under review at NeurIPS 2025.

UCFE: A User-Centric Financial Expertise Benchmark for Large Language Models [online]
Yuzhe Yang*, Yifei Zhang*, Yan Hu*, Yilin Guo, Ruoli Gan, Yueru He, Mingcong Lei, Xiao Zhang, Haining Wang, Qianqian Xie, Jimin Huang, Honghai Yu, Benyou Wang
In Findings of the Association for Computational Linguistics: NAACL 2025, pages 5429–5448.

Open-FinLLMs: Open Multimodal Large Language Models for Financial Applications [online]
Jimin Huang, Mengxi Xiao, Dong Li, Zihao Jiang, Yuzhe Yang, Yifei Zhang, Lingfei Qian, Yan Wang, Xueqing Peng, et al.
arXiv preprint arXiv:2408.11878 (2024).

FAST-CA: Fusion-based Adaptive Spatial-Temporal Learning with Coupled Attention for airport network delay propagation prediction [online]
Chi Li, Xixian Qi, Yuzhe Yang, Zhuo Zeng,Lianmin Zhang, Jianfeng Mao
Information Fusion 107 (2024): 102326.

FedDTPT: Federated Discrete and Transferable Prompt Tuning for Black-Box Large Language Models [online]
Jiaqi Wu, Simin Chen, Yuzhe Yang, Yijiang Li, Shiyue Hou, Rui Jing, Zehua Wang, Wei Chen, Zijian Tian
arXiv preprint arXiv:2411.00985 (2024).

FDPT: Federated Discrete Prompt Tuning for Black-Box Visual-Language Models
Jiaqi Wu, Simin Chen, Jing Tang, Yuzhe Yang, Yiming Chen, Lixu Wang, Song Lin, Zehua Wang, Wei Chen, Zijian Tian
Under review at ICCV 2025.

Feature Extraction and Steering for Enhanced Chain-of-Thought Reasoning in Language Models
Zihao Li, Xu Wang, Yuzhe Yang, Ziyu Yao, Haoyi Xiong, Mengnan Du
Under review at EMNLP 2025.

(* Equal Contribution,)

Research Experience

Large-scale Social Simulation Agent — Oct. 2024 – Jan. 2025
Advised by Prof. Benyou Wang & Prof. Honghai Yu — *CUHK-Shenzhen & Nanjing University*

- Developed the TwinMarket framework where LLM agents simulate large-scale human investor behaviors to validate economic principles and market theories, leveraging the TwinMarket framework to model individual decision-making and social interactions.
- Built scalable simulations to evaluate how collective trading behavior impacts broader market outcomes, successfully replicating key market phenomena such as volatility clustering and fat-tailed return distributions.
- This work won the Best Paper Award at *ICLR Workshop 2025*

Financial Multimodal Large Language Model May. 2024 – Oct. 2024

Advised by Prof. Benyou Wang & Jimin Huang *CUHK-Shenzhen & TheFinAI*

- Leaded the multimodal extension of LLM
- Developed a multimodal financial benchmark dataset for LLM training and evaluation
- Multimodal instruction finetuning for LLM, include text, image (chart & tabular) and numerics data
- Align multimodal LLM with financial data and real-world scenarios to improve model performance
- Released FinLLaVA-8B: Achieved MMMU (Overall) score of 36.3 and MMMU (Business) score of 30.7
- Constructed a purely text-based multi-turn dialogue benchmark to evaluate the performance of LLMs in real-world financial applications using a user simulator; this work had accepted to *NAACL Findings 2025*; #1 Paper of the day on Hugging face

Flight Delay Propagation Modeling Aug. 2023 – Sep. 2024

Advised by Prof. Jianfeng Mao *CUHK-Shenzhen*

- Developed a GNN framework integrating dynamic and adaptive graph learning with coupled attention mechanisms to address complex spatial-temporal dependencies in airport delay propagation; this work had published in *Information Fusion*
- Enhanced the SIS epidemiological model by incorporating adaptive graph learning to simulate and predict epidemic transmission dynamics in airport networks; this work is in preparation, to be submitted to *Transportation Research Part B: Methodological*
- Leverages Neural ODE networks to improve flight delay prediction by developing a continuous graph model that enhances interpretability, reduces training time, and addresses challenges like irregular time sampling and missing data

Projects

Quant-GPT: Money is All You Need [online] | *PyTorch, Transformers, ChromaDB* Mar. 2024 – Apr. 2024

- Final project for the PhD course CSC6052, a multi-agent system for A-share market investment decisions
- Fine-tuned an LLM, integrating it with sentiment analysis and real-world market data.
- Utilized RAG and multi-agent systems to dynamically access and synthesize relevant financial news, enhancing the model's ability to forecast market trends and returns
- Results achieved: Sharpe Ratio: 0.40, Annualized Return: 7.26%, Max Drawdown: 13.61%

Travel Insurance Recommendation AI System [online] | *PyTorch, LangChain* Jan. 2024 – Apr. 2024

- Developed an AI system to predict flight delays and recommend personalized travel insurance
- Fine-tuned the LLM using an insurance corpus to improve domain-specific question-answering capabilities, achieving an 83% accuracy in identifying user intent
- Utilized deep learning and LLM agents for accurate delay predictions and customer sentiment assessment

Flight Information System [online] | *Python, LangChain, SQL, Flask* Mar. 2024 – Apr. 2024

- Developed database system to optimize airline management, including passenger bookings and flight logistics
- Delivered a functional database with a user-friendly web interface
- Integrated LLM to enhance database architecture and query generation

Technical Skills

Languages: Python, C/C++
Developer Tools: Git, Docker, Linux, Slurm
Libraries: PyTorch, Transformers

Awards

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

Service

Reviewer: IJCAI 2025, ICLR 2025 Workshop

Latest Version
Last updated on May, 2025