Enpei Zhang

✉ enpei.zhang.gr@dartmouth.edu | 🌐 Website | ⑧ Google Scholar

RESEARCH INTERESTS

My research interests center on **AI agents**, **Reinforcement Learning**, and **large language models**, with an emphasis on how agent interactions and workflow structures shape performance, efficiency, and reliability. My recent research focuses on principled methods for automatic **agent workflow orchestration**, grounded in game theory and scheduling theory to improve efficiency and incentive alignment.

EDUCATION

• **Dartmouth College**   *Sep 2024 - Present*
Ph.D. student in Computer Science   Hanover, NH, USA
  ◦ **Advisor**: Prof. Yujun Yan

• **Shanghai Jiaotong University**   *Sep 2020 - Jun 2024*
Bachelor of Science (Zhiyuan Honor Class), Information Engineering ; GPA: 89/100   Shanghai, China
  ◦ **Advisor**: Prof. Siheng Chen

SELECTED PUBLICATIONS

[Nature Communications] Incentivizing inclusive contributions in model sharing markets   Article Link
  Enpei Zhang, Jingyi Chai, Rui Ye, Yanfeng Wang, Siheng Chen
  Proposed iPFL, an incentive-compatible personalized federated learning framework that formulates federated training as a multi-agent market, enabling heterogeneous data holders to collaboratively train personalized models with formal guarantees of individual rationality and incentive compatibility, while achieving strong performance and economic utility across diverse training tasks including classification and LLM SFT.
  🏷 Multi-Agent System 🏷 Game Theory 🏷 Federated Learning

[EMNLP 2025] Judging with Many Minds: Do More Perspectives Mean Less Prejudice?   Article Link
  Chiyu Ma*, **Enpei Zhang***, Yilun Zhao, Wenjun Liu, Yaning Jia, Peijun Qing, Lin Shi,
  Arman Cohan, Yujun Yan, Soroush Vosoughi
  Conducted a systematic evaluation of bias behaviors in multi-agent LLM-as-Judge frameworks, modeling evaluation as an interactive multi-agent system rather than a single-judge process. The results revealed the counterintuitive finding that multi-agent debate can exacerbate, rather than reduce, four distinct biases over multiple debate rounds.
  🏷 Multi-Agent System 🏷 LLM Bias

[ICLR 2026 Oral] Seeing Through the Brain: New Insights from Decoding Visual Stimuli with fMRI   Article Link
  Zheng Huang*, **Enpei Zhang***, Yinghao Cai, Weikang Qiu, Carl Yang, Elynn Chen, Xiang Zhang,
  Rex Ying, Dawei Zhou, Yujun Yan
  Demonstrated that fMRI signals align better with structured text representations than vision-model representations. Proposed PRISM, a text-based framework that improves fMRI-to-image reconstruction via object-centric diffusion.
  🏷 AI for Science 🏷 Representation Learning

ONGOING WORKS

[In Submission] Multi-Agent Workflow Generation with Makespan Minimization
  First author; manuscript under review.
  Formulated latency-aware multi-agent workflow generation as a principled optimization problem that minimizes execution makespan while preserving downstream task performance. Proposed a theoretically grounded framework that eliminates critical-path bottlenecks via structure-level workflow optimization, enabling both de novo workflow generation and refinement of existing agent pipelines.
  🏷 Multi-Agent System 🏷 LLM Agent 🏷 Scheduling Theory

[In Submission] Optimistic Transfer under Task Shift via Bellman Alignment
  Second author; manuscript under review.
  Studied online transfer reinforcement learning under episodic MDPs with task shift, identifying one-step Bellman alignment as the correct abstraction for statistically sound transfer. Proposed an operator-level correction that reweights the value from source task to target task, reducing task mismatch by a fixed one-step bias correction. The resulting two-stage reweighted Q learning framework decouples variance reduction from bias correction and admits regret bounds scaling with task-shift complexity rather than target MDP complexity.
  🏷 Reinforcement Learning 🏷 Transfer Learning

EXPERIENCE

- **Cooperative Medianet Innovation Center, Shanghai Jiaotong University** *Sep 2022 - Jun 2024*
Research Assistant Shanghai, China
Federated Learning and Incentive Mechanism Design for Multi-Agent Collaboration
Article | Advised by Prof. Siheng Chen. Investigated game-theoretic mechanism design in multi-agent collaboration. Specifically, we focused on incentive problems in cross-institution federated learning and proposed an incentive-compatible mechanism that enables collaborative training among strategic data holders. The work was published in *Nature Communications*.

TEACHING & SERVICES

Graduate Teaching Assistant
- COSC74/274 (Machine Learning), Dartmouth College, Winter 2025.
- COSC89/189 (Network Science and Complex Systems), Dartmouth College, Fall 2024.

Reviewer
- ICML (International Conference on Machine Learning) 2026
- ACM Transactions on Knowledge Discovery

HONORS & AWARDS

- **Dartmouth Fellowship,** Dartmouth College *2024-Present*
- **Outstanding Graduate Award,** Shanghai Jiaotong University *2024*
- **Zhiyuan Honor Scholarship,** Shanghai Jiaotong University *2020–2023*