Chenmien Tan

Email: chenmientan@outlook.com ⋄ Phone: 185 1280 4143
Homepage | Google Scholar | GitHub | Kaggle

PROFESSIONAL EXPERIENCE

- **Alibaba International Digital Commerce, Alibaba Group** 2025/02–present
  Algorithm Engineer
  - Develop rubric-based content quality and fine-grained citation faithfulness evaluation to monitor model performance, filter data for SFT, and provide reward signal for RL
  - Synthesize user queries, simulated search results, and responses by few-shot prompting and refining distillation data, respectively

EDUCATION

- **University of Edinburgh** 2023/09–2024/09
  M.Res. in Language Processing, ***only admit 3 students****, Supervisor: Shay Cohen*
- **University of Nottingham Ningbo China** 2018/09–2022/06
  B.Sc. in Applied Mathematics

PROJECTS

- **RL2: Ray Less Reinforcement Learning** [Code] [Blog]
  Lead developer, ***>900 GitHub stars****, liked by Andrej Karpathy and reposted by Sheng Ying*
  - Support SFT, RM, DPO, and PPO, where the performance and efficiency are comparable with other popular LLM post-training libraries
  - Support 3D parallelism (DP/CP/TP) with FSDP backend and 5D parallelism (DP/CP/PP/TP/EP) with Megatron backend, enabling scalability to model scale and context length
  - Support partial rollout and async inference engine to eliminate GPU idle caused by length imbalance and environment latency, respectively
  - Balance load of training engines, inference engines, and environments by balanced sequence packing, router forwarding, and resource pooling, respectively
  - Scale agentic environments by integrating GEM and support non-concat state transition by cross-sequence advantage computation

SELECTED PUBLICATIONS

- **GEM: A Gym for Agentic LLMs** [Paper] [Code]
  Co-author, >400 GitHub stars, Outstanding in NeurIPS'25 SEA workshop
  - Provide unified OpenAI Gym-like API for various environments (*e.g.*, coding and question-answering) and tools (*e.g.*, Python executor and search engine)
  - Propose REINFORCE-like algorithm with credit assignment, which enjoys superior performance than other popular LLM RL algorithms in multi-turn scenario
- **Massive Editing for Large Language Models via Meta Learning** [Paper] [Code]
  First author, ***>130 citations, top 10% cited paper in ICLR'24***
  - Improve the scalability of LLM editing hyper-network by formulating the parameter shift aggregation as a least square problem to mitigate cancellation effect

  - Decrease memory peak of editing LLM through delineating computation between hyper-network and LLM

- **CVaR-Regret Bounds for Multi-Armed Bandits** [Paper]
  First author, ACML'22
  - Propose algorithms enjoying nearly optimal performance on the tail risk of stochastic and adversarial multi-armed bandit problems with theoretical guarantees
  - Reveal the characteristics of implicit exploration for achieving trade-off between the expected regret and variability of regret

COMPETITIONS

- Learning Equality – Curriculum Recommendations 2023/01–2023/03
  Ranking: 17/1507 = 1.6%
- Google AI4Code – Understand Code in Python Notebooks [Blog] 2022/06–2022/08
  Ranking: 25/1135 = 2.2%
- H&M Personalized Fashion Recommendations [Code] 2022/03–2022/05
  Ranking: 45/2952 = 1.5%

ACADEMIC SERVICES

- Reviewer: ICML'24-25, NeurIPS'25, ICLR'25-26
- Teaching assistant: NLU+ 23-24, MLP 23-24