Yucheng Pan

+86 193 7248 5548 ◇ yuchengpan707@gmail.com ◇ Personal Homepage

EDUCATION

B.S. Mathematics and Physics, Tsinghua University Sep. 2021 - Jun. 2025 (expected)
GPA: 3.874/4.000 Beijing, China

RESEARCH INTERESTS

Primary Large Language Model Agent (LLM Agent), Reinforcement Learning
Supplementary AI Safety, Multi-Agent Systems, Lifelong Learning, Bio-Inspired AI

RESEARCH EXPERIENCE

Research Internship, Princeton University (Remote) May. 2024 - Present
Department of Electrical and Computer Engineering
Mentor: Prof. Mengdi Wang (Princeton ECE), Prof. Huazheng Wang (Oregon State EECS)

· ***LLM Agents Can Deceive and Be Constrained by Social Norms and Payoff Allocations***

- **Goal**: Investigate how LLM agents exhibit deceptive behaviors in multi-agent games and design mechanisms to enforce compliance with social norms and fair payoff distributions.
- Co-led the literature review and ideation process, synthesizing insights to define key research questions.
- Designed the overall framework, implemented algorithms, and conducted extensive experiments with comprehensive analysis of results.
- Contributed to a manuscript submitted to *Nature Computational Science*.

· ***LLM Agent Data Augmentation***

- **Goal**: Enhance the generalization capabilities of LLM agents through innovative data augmentation techniques.
- Played a key role in literature review, preliminary experimentation, and identifying gaps in current methodologies.
- Proposed novel methods to redesign agent environments for fine-tuning, aligning with project objectives.

Research Internship, University of North Carolina at Chapel Hill (Remote) Feb. 2024 - Oct. 2024
Department of Statistics and Operations Research
Mentor: Prof. Yao Li (UNC Chapel Hill STOR), Prof. Minhao Cheng (Penn State IST)

· ***Orthogonal Audio Watermarking with Multi-Embedding***

- **Goal**: Develop a novel audio watermarking method that enables embedding multiple watermarks from different sources into a single audio file to improve robustness and traceability.
- Led the literature review and idea development.
- Redesigned the architecture and fine-tuning pipeline of AudioSeal for multi-embedding support.
- Conducted rigorous experiments and comprehensive analysis to validate the proposed approach.

Undergraduate Research Assistant, Tsinghua University Sep. 2023 - Jul. 2024
Department of Statistics and Data Science
Mentor: Prof. Ke Deng

· ***Enhancing Precision in Isotope Nuclear Radius Estimation through Statistical Analysis***

- **Goal**: Improve the precision of isotope nuclear radius estimation using advanced statistical analysis of laser spectroscopy data.
- Led the literature review, data collection, and pre-processing to establish a robust foundation for the study.
- Developed statistical models, implemented statistical computing algorithms to enhance the precision and validated the results.

· ***Nucleotide Sequences Prediction and Design Based on Language Models***

- **Goal**: Utilize BERT models to predict nucleotide sequences and explore their applications in artificial protein design.
- Led the implementation of predictive models and conducted extensive experiments to verify the feasibility of the approach.

SCIENTIFIC TALKS

Oral Presentation Jul. 2024
Topics on Frontiers of Cross-Sciences, Beijing
· Enhancing Isotope Charge Radius Measurement Precision with Statistical Analysis

Oral Presentation Dec. 2023
Tsinghua Text Analysis Symposium, Beijing
· PLMs as Meta-function: Learning In-context Learning for Named Entity Recognition

SCHOLARSHIPS AND AWARDS

Academic Excellence Scholarship, Tsinghua University 2024
Progress in Academic Performance Scholarship, Tsinghua University 2022
Outstanding Innovation Scholarship, Philip K.H. Wong Foundation 2022

SELECTED COURSE PROJECTS

Large Language Models and Alignment Sep. 2024 - Present
· Pre-training, instruction fine-tuning, and RLHF on LLMs, with a focus on CUDA/DPU programming.

Deep Reinforcement Learning Mar. 2024 - Jun. 2024
· Conducted literature reviews, designed experiments, and developed algorithms to improve offline RL performance under limited data scenarios.
· Delivered a project paper and oral presentation.

Machine Learning and Big-data Nov. 2023 - Dec. 2023
· Designed and implemented neural networks (ANN, RNN, CNN) to predict autonomous underwater vehicle health.