Peng Wang

Linkedin: https://www.linkedin.com/in/pengw96/
Website: https://peter-peng-w.github.io/

Email : pw7nc@virginia.edu
Mobile : +1-434-328-9834

Education

- **University of Virginia** — Charlottesville, VA
  Ph.D. Student in Computer Science, Advisor: Jing Yang and Cong Shen — *Aug. 2022 – Present*
- **University of Virginia** — Charlottesville, VA
  Master of Science in Computer Science, Advisor: Hongning Wang — *Aug. 2019 – Dec. 2021*
- **Tsinghua University** — Beijing, China
  Bachelor of Engineering in Computer Science and Technology — *Sept. 2014 – Jun. 2018*

Publications

[1] R. Liu, **P. Wang**, D. Li, C. Shen, and J. Yang, "A shared low-rank adaptation approach to personalized rlhf," *International Conference on Artificial Intelligence and Statistics*, 2025.

[2] S. Wang*, **P. Wang***, T. Zhou, Y. Dong, Z. Tan, and J. Li, "Ceb: Compositional evaluation benchmark for fairness in large language models," *The Thirteenth International Conference on Learning Representations*, 2025, Spotlight Paper.

[3] S. Wang, **P. Wang**, T. Zhou, *et al.*, "On demonstration selection for improving fairness in language models," in *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, *Workshop on Socially Responsible Language Modelling Research*, Spotlight Paper, 2024.

[4] **P. Wang**, R. Cai, and H. Wang, "Graph-based extractive explainer for recommendations," in *Proceedings of the ACM Web Conference 2022*, 2022, pp. 2163–2171.

RESEARCH INTEREST

- My research interests span various topics in machine learning, including information retrieval, reinforcement learning, and trustworthy AI. Recently, I have been particularly interested in exploring **LLM alignment** techniques to improve the faithfulness of generated responses and **improve models' reasoning abilities**. Furthermore, I am interested in the **trustworthiness of LLMs**, including (but not limited to) their robustness against malicious attacks during instruction tuning and fairness issues in both training-free evaluation (e.g., through in-context learning) and alignment tuning.

Skills Summary

- **Programming Languages**: Adept at Python, C/C++, familiar with Linux, Java, R, SQL
- **Machine Learning**: Adept at PyTorch, familiar with TensorFlow

Technical Research

- **LLM Reasoning** — Charlottesville, USA
  Research Assistant, Directed by Prof. Jing Yang and Prof. Cong Shen, University of Virginia — *Sep. 2024 - Present*
  - Incorporating partially observable Markov decision processes (POMDPs) into the stepwise sampling-based reasoning framework to shorten the reasoning trajectory and denoise erroneous reasoning steps.
  - Distilled state representations from large language models (e.g., GPT-4o) by leveraging integrated forward planning and retrospective self-reflection processes, thereby promoting downstream decision-making and model interpretability.
  - Leveraging reinforcement learning (e.g., PPO and GRPO) to optimize state representations derived from distilled structured content, further enhancing model reasoning capabilities and overall performance.
- **Alignment of LLM** — Charlottesville, USA
  Research Assistant, Directed by Prof. Hongning Wang, University of Virginia/Tsinghua University — *Sep. 2023 - Present*
  - Introduced Reward/Advantage-weighted Regression to promote model's alignment performance during both SFT and DPO.
  - Investigating data selection and generation methods that integrate trajectory rewards to enhance multi-step reasoning in formal mathematical proof generation.
- **Fairness in LLM** — Charlottesville, USA
  Research Assistant, Directed by Prof. Yangfeng Ji and Prof. Jundong Li, University of Virginia — *Jan. 2024 - Present*
  - Developed a synthesized benchmark to assess LLMs' zero-shot and few-shot fairness across various tasks, including stereotype recognition/classification, toxic content generation, and decision-making based on sensitive attributes.
  - Proposed multi-stage clustering strategies to adaptively select in-context demonstrations, improving LLMs' group fairness in decision-making tasks.

- **Explainable Recommendation (XRec)** Charlottesville, USA
  Research Assistant, Directed by Prof. Hongning Wang, University of Virginia *Sep. 2020 - May. 2023*
  - Reimplemented baseline models including NRT and Att2Seq and evaluated them on datasets including Yelp and TripAdvisor.
  - Proposed to use graph structure to model the relationship between user, item, attributes and candidate explanations.
  - Leveraged on Graph Attention Network to predict the relevance score of each candidate sentences to form explanations.
  - Conducted data poisoning attacks on matrix-based and neural network-based XRec methods to investigate their robustness.
- **Continual Reinforcement Learning** Los Angeles, USA
  Research Assistant, Directed by Prof. Yan Liu, University of Southern California *Jul. 2018 - Oct. 2018*
  - Reproduced DQN, Double DQN, Duel DQN and Prioritized Experience Replay and evaluated them on Atari games.
  - Implemented various unsupervised representation learning methods to improve the training speed of the current DQN method.
  - Combined DQN with a novel expandable neural network structure to achieve continual RL.

Work Experience

- **Zhipu AI** Beijing, China
  Machine Learning Intern, RLHF Group *Jun. 2024 - Aug. 2024*
  - Worked on LLM post-training for Automatic Theorem Proving in Lean.
  - Implemented multiple search strategies including whole-proof sampling, per-step tactic generation via best-first search, and MCTS, which were then used to synthesize theorem proofs to scale up the supervised fine-tuning dataset.

Service

- Reviewer of ACM TIST, ICLR'25, subreviewers of KDD'22, WWW'23, AAAI'24