

# Qingfeng Lan

## Curriculum Vitae



**Research** Enhancing large language models with reinforcement learning (RL).

**Interest** Improving RL efficiency by reducing forgetting and maintaining plasticity.  
Understanding the inner workings of deep neural networks.

### Education

20.09 - 25.09 Doctor of Philosophy in Computing Science, **University of Alberta**, Canada.

Thesis Toward Efficient Reinforcement Learning Under Non-Stationarity. [\[Link\]](#)

Supervisor [A. Rupam Mahmood](#)

18.09 - 20.08 Master of Science in Computing Science, **University of Alberta**, Canada.

Thesis Predictive Representation Learning for Language Modeling. [\[Link\]](#)

Supervisor [Alona Fyshe](#)

14.09 - 18.07 Bachelor of Engineering in Computer Science and Technology, **University of Chinese Academy of Sciences**, China.

Thesis A Deep Top-K Relevance Matching Model for Ad-hoc Retrieval. [\[Link\]](#)

Advisor [Yanyan Lan](#) (thesis advisor), [Guojie Li](#) (tutor)

17.10 - 18.03 Visiting Non-Matriculated Programme, **University of Oxford**, England.

Tutor [Leslie Ann Goldberg](#)

### Employment

25.10 - Present Researcher, **Qwen Team, Alibaba Group**, Beijing, China.

Project Enhancing large language models with reinforcement learning.

24.11 - 25.04 Research Intern, **Huawei Noah's Ark Lab**, Edmonton, Canada.

Collaborator [Chao Gao](#)

Project Enhancing large language models with reinforcement learning.

24.06 - 24.10 Research Intern, **Meta Reality Lab**, California, United States.

Collaborator [Rohan Chitnis](#), [Alborz Geramifard](#), [Ta-Chu Kao](#), [Jorge Menendez](#)

Project Improving Next-generation Wearables through Reinforcement Learning. [\[Link\]](#)

o Built a real-time online reinforcement learning training pipeline from scratch.

o Improved the cursor control policy of the wearables with reinforcement learning.

22.07 - 23.01 Research Intern, **Sea AI Lab**, Singapore.

Collaborator [Zhongwen Xu](#), [Shuicheng Yan](#)

Project Learning to Optimize for Reinforcement Learning. [\[Link\]](#)

- Applied meta-learning to learn an optimizer for reinforcement learning tasks.
  - Proposed the first learned optimizer for reinforcement learning that is stable to train and generalizes to unseen tasks. Paper accepted at RLC 2024.
- 22.01 - 22.06 Research Intern, **Huawei Noah's Ark Lab**, Edmonton, Canada.  
 Collaborator [Yangchen Pan](#), [Jun Luo](#)  
 Project Memory-efficient Reinforcement Learning with Value-based Knowledge Consolidation. [\[Link\]](#)
  - Demonstrated that catastrophic forgetting exists even in single-task reinforcement learning, resulting in low learning efficiency.
  - Reduced the replay buffer size significantly by mitigating forgetting with value-based knowledge consolidation. Paper accepted at TMLR.
- 17.07 - 18.04 Research Assistant, **Key Laboratory of Network Data Science and Technology, Chinese Academy of Sciences**, Beijing, China.  
 Collaborator [Yixing Fan](#), [Yanyan Lan](#), [Jiafeng Guo](#)  
 Project A Deep Top-K Relevance Matching Model for Ad-hoc Retrieval. [\[Link\]](#)
  - Proposed a deep relevance matching model for ad-hoc retrieval problem by applying the top-k pooling and a term gating network.
  - Outperformed SOTA models on two representative benchmark datasets. Paper accepted at CCIR 2018.

## Publications

\*: Equal contribution

### Conference and Journal Articles

- Nature-2024 **Loss of Plasticity in Deep Continual Learning.**  
 Shubhang Dohare, J. Fernando Hernandez-Garcia, **Qingfeng Lan**, Parash Rahman, A. Rupam Mahmood, Richard S. Sutton. *Nature*, 2024. Article. [\[Link\]](#)
- RLC-2024 **Learning to Optimize for Reinforcement Learning.**  
**Qingfeng Lan**, A. Rupam Mahmood, Shuicheng Yan, Zhongwen Xu. *Reinforcement Learning Conference*, 2024. Oral. [\[Link\]](#)
- RLC-2024 **More Efficient Randomized Exploration for Reinforcement Learning via Approximate Sampling.**  
 Haque Ishfaq, Yixin Tan, Yu Yang, **Qingfeng Lan**, Jianfeng Lu, A. Rupam Mahmood, Doina Precup, Pan Xu. *Reinforcement Learning Conference*, 2024. Oral. [\[Link\]](#)
- RLC-2024 **Weight Clipping for Deep Continual and Reinforcement Learning.**  
 Mohamed Elsayed, **Qingfeng Lan**, Clare Lyle, A. Rupam Mahmood. *Reinforcement Learning Conference*, 2024. Oral. [\[Link\]](#)
- ICLR-2024 **Provable and Practical: Efficient Exploration in Reinforcement Learning via Langevin Monte Carlo.**  
 Haque Ishfaq\*, **Qingfeng Lan**\*, Pan Xu, A. Rupam Mahmood, Doina Precup, Anima Anandkumar, Kamyar Azizzadenesheli. *International Conference on Learning Representations*, 2024. Poster. [\[Link\]](#)

- TMLR-2023 **Memory-efficient Reinforcement Learning with Value-based Knowledge Consolidation.**  
**Qingfeng Lan**, Yangchen Pan, Jun Luo, A. Rupam Mahmood. *Transactions on Machine Learning Research*, 2023. **CoLLAs certification.** [\[Link\]](#)
- AISTATS-2022 **Model-free Policy Learning with Reward Gradients.**  
**Qingfeng Lan**, Samuele Tosatto, Homayoon Farrahi, A. Rupam Mahmood. *International Conference on Artificial Intelligence and Statistics*, 2022. **Poster.** [\[Link\]](#)
- ICLR-2020 **Maxmin Q-learning: Controlling the Estimation Bias of Q-learning.**  
**Qingfeng Lan**, Yangchen Pan, Alona Fyshe, Martha White. *International Conference on Learning Representations*, 2020. **Poster.** [\[Link\]](#)
- CCIR-2018 **A Deep Top-K Relevance Matching Model for Ad-hoc Retrieval.**  
Zhou Yang, **Qingfeng Lan**, Jiafeng Guo, Yixing Fan, Xiaofei Zhu, Yanyan Lan and Yue Wang, Xueqi Cheng. *China Conference on Information Retrieval*, 2018. **Best Paper Award Candidate.** [\[Link\]](#)  
[Workshop and Non-Refereed Articles](#)
- arXiv-2025 **Dynamic Speculative Agent Planning.**  
Yilin Guan, **Qingfeng Lan**, Sun Fei, Dujian Ding, Devang Acharya, Chi Wang, William Yang Wang, Wenyue Hua. *arXiv*, 2025. [\[Link\]](#)
- arXiv-2025 **Efficient Reinforcement Learning by Reducing Forgetting with Elephant Activation Functions.**  
**Qingfeng Lan**, Gautham Vasan, A. Rupam Mahmood. *arXiv*, 2025. [\[Link\]](#)
- ICML-2023 **Elephant Neural Networks: Born to Be a Continual Learner.**  
**Qingfeng Lan**, A. Rupam Mahmood. *ICML Workshop on High-dimensional Learning Dynamics*, 2023. **Poster.** [\[Link\]](#)
- EWRL-2023 **Overcoming Policy Collapse in Deep Reinforcement Learning.**  
Shibhansh Dohare, **Qingfeng Lan**, A. Rupam Mahmood. *European Workshop on Reinforcement Learning*, 2023. **Poster.** [\[Link\]](#)
- arXiv-2021 **Variational Quantum Soft Actor-Critic.**  
**Qingfeng Lan**. *Quantum Computing Course Project*, 2021. [\[Link\]](#)

---

## Open-Source Code

[Jaxplorer](#).

A Jax reinforcement learning framework for exploring new ideas.

[Optim4RL](#).

A Jax framework of learning to optimize for reinforcement learning.

[Explorer](#).

A PyTorch reinforcement learning framework for exploring new ideas.

[Gym Games](#).

A collection of Gymnasium compatible games for reinforcement learning.

[Quantum Explorer](#).

A quantum reinforcement learning framework based on PyTorch and PennyLane.

[Loss of Plasticity](#).

The implementation of continual backpropagation which maintains network plasticity.