Lihong Li

Redmond, WA 98052 • lihongli.cs@gmail.com • (646) 812-1119

My core research interest is in machine learning for interactive, agentic systems that maximizes a reward function by taking actions, including **large language models**, **reinforcement learning**, **contextual bandits**, and related areas. My work has been applied to **conversational systems, recommendation, advertising, Web search** in the industry.

[[Homepage](#) | [Full Resume](#) | [Google Scholar](#) | [LinkedIn](#)]

PROFESSIONAL EXPERIENCE

META   **Bellevue, WA**
AI Research Scientist   **2025-present**
- Working at the intersection of large language models, reinforcement learning, and recommendation systems.

AMAZON   **Seattle, WA**
Senior Principal Scientist   **2020-2025**
- Science lead for reinforcement learning based LLM post-training of Amazon's Rufus shopping foundation models, to elicit reasoning and agentic capabilities, and to align with shopping customer preferences and Amazon Stores policies.
- Science lead for one of the earliest GenAI/LLM adoptions in Sponsored Products, to allow advertisers to reach audiences with selected targeting themes, and with optimized campaign performance.
- Science lead for reinforcement learning based solutions to measure long-term impacts on customer behavior, based on short-term experimentation data (such as in A/B tests). The measurement is critical in experimentation-based decision making, to optimize the tradeoff of organic and sponsored results.
- Designed and initiated an Early Career Scientist program in Sponsored Ads, which directly inspired and influenced the company-wide Postdoctoral Scientist program 1 year after.

GOOGLE   **Kirkland, WA**
Research Scientist   **2017-2020**
- Research and development in offline/off-policy reinforcement learning, to predict and optimize long-term rewards based on offline data, targeting various applications such as healthcare.
- One of the founding members of Google AI China Center, the first such center in Asia. In addition to conducting AI research, also heavily involved in hiring and engagement with local academic communities.

MICROSOFT   **Redmond, WA**
Researcher 2012-2017
- Research in contextual bandits and reinforcement learning algorithms and theory.
- Deploy contextual bandits / RL techniques in close collaboration with product teams like Bing.
- Research in deep reinforcement learning, especially in conversational systems (chatbots).

YAHOO!   **Santa Clara, CA**
Research Scientist   **2009-2012**

- Research and development of contextual bandit solutions in Yahoo!'s personalized news recommendation. Among the first to apply RL in a critical business domain in the industry. Developed fundamental algorithms, data collection strategies, and evaluation protocols.
- Research in large-scale machine learning, with contributions to Vowpal Wabbit open source.

EDUCATION

RUTGERS UNIVERSITY
Ph.D. Computer Science

New Brunswick, NJ
2005-2009

UNIVERSITY OF ALBERTA
M.S. Computing Science

Edmonton, AB, Canada
2002-2004

TSINGHUA UNIVERSITY
B.E. Computer Science and Technology

Beijing, China
1998-2002

SELECTED AWARDS

- Seoul Test of Time Award, 2023
- Yahoo! Super Star Team Award (highest team achievement award in the company), 2011
- AISTATS Notable Paper Award, 2011
- WSDM Best Paper Award, 2011
- ICML Best Student Paper Award, 2008

SELECTED PUBLICATIONS

- Dualdice: Behavior-agnostic estimation of discounted stationary distribution corrections (NeurIPS 2019)
- Breaking the curse of horizon: Infinite-horizon off-policy estimation (NeurIPS 2018)
- SBEED: Convergent reinforcement learning with nonlinear function approximation (ICML 2018)
- Neural approaches to conversational AI (NOW Publishers, 2018)
- Doubly robust off-policy value evaluation for reinforcement learning (ICML 2016)
- Taming the monster: A fast and simple algorithm for contextual bandits (2014)
- An empirical evaluation of Thompson sampling (NeurIPS 2011)
- Contextual bandits with linear payoff functions (AISTATS 2011)
- Unbiased offline evaluation of contextual-bandit-based news article recommendation algorithms (WSDM 2011)
- A contextual-bandit approach to personalized news article recommendation (WWW 2010)
- Knows what it knows: a framework for self-aware learning (ICML 2008)
- Sparse online learning via truncated gradient (NeurIPS 2008)
- PAC model-free reinforcement learning (ICML 2006)

ADDITIONAL INFORMATION

- Internship at Google (2006), Yahoo! Labs (2007), and AT&T Shannon Lab (2008)
- Regularly serves as Senior Area Chairs at leading AI venues (ICLR, ICML, NeurIPS)
- Associate, acting, and guest editor for IEEE TPAMI, TMLR, and MLJ