Xiaoyu Lu

luxiaoyu644@gmail.com
Personal Website: https://xiaoyulu2022.github.io/
Google Scholar: https://scholar.google.com/citations?hl=en&user=ELqAe3MAAAAJ

About Me

I have been working at Amazon for $\sim$ 5 years, where I have research experience and developed multiple innovative initiatives using statistical and machine learning modelling which have been deployed in production. I am specialized in the field of statistical science, including topics and applications in probabilistic inference, machine learning, explainable artificial intelligence (XAI), causal inference, deep generative modelling. Prior to Amazon, I completed my PhD in Statistical Science at University of Oxford, supervised by Prof. Yee Whye Teh in the Machine Learning group at the Department of Statistics, with research experience in Gaussian Processes, reinforcement learning, causal inference, Bayesian optimisation and deep generative models. Before my PhD, I did my undergraduate MMath degree in Mathematics and Statistics at University of Oxford.

Work Experience

12/2023 - Present, Applied Scientist, Amazon, US

- Research and develop machine learning or statistical models for supply chain systems.
- Invent and build the Emulation technology as an experimentation platform for inventory planning, leveraging machine learning and statistical models.
- Invented and developed a data-driven, statistical model to better predict vendors' behavior, which has important business impact on Amazon's supply chain systems. The model has been launched in shadow mode in O4 2023 and will be launched in production in Q1 2024.

04/2019 - 12/2023, Machine Learning Scientist, Amazon, UK

- Delivered a research project on explainable AI. My work on this topic has been published at the International Conference on Machine Learning (ICML) 2022.
- Invented and built explainable emulators for complex Amazon supply chain systems, which is the technology behind the Attribution service that explains evolution of key Amazon inventory prediction metrics. The technology was deployed in production since Q4 2022.
- Delivered multiple research projects on the topics and applications of Bayesian optimization, causal inference, reinforcement learning, with publications at Amazon internal and prominent external conferences. Participated in the research community through judging the work of others, giving talks and being a committee member in peer-reviewed internal and external conferences.
- Outside my role at Amazon, I worked on the "Daisee" project - a statistical inference method based on adaptive importance sampling which trades off exploration versus exploitation. My work has been accepted by the Scandinavian Journal of Statistics, 2023.

Education

10/2014 - 06/2019, PhD in Statistical Science, University of Oxford, UK

- I obtained my PhD degree working in the field of probabilistic machine learning with Prof. Yee Whye Teh. I have research experience in probabilistic generative models, Bayesian inference, Gaussian Process, Markov Chain Monte Carlo (MCMC), causal inference and reinforcement learning, with publications at prominent statistics and machine learning conferences.

10/2010 - 06/2014, MMath in Mathematics and Statistics, University of Oxford, UK

- First Class degree (ranked top 1st in both Bachelor and Master).
- Scored 83% in fourth year thesis on Recommender System for movie recommendations.

Selected Publications and Papers

Research Profile available at Google Scholar. Citation:187.
Amazon Science: https://www.amazon.science/author/xiaoyu-lu

- **X. Lu**, T. Rainforth, T, Y. W. Teh, Daisee: Adaptive Importance Sampling by Balancing Exploration and Exploitation, in Scandinavian Journal of Statistics, 2023.

- **X. Lu**, A. Boukouvalas, J. Hensman, Additive Gaussian Processes Revisited, in International Conference on Machine Learning (ICML), 2022.
- V. Aglietti, **X. Lu**, A. Paleyes, J. González, Causal Bayesian Optimization, in Artificial Intelligence and Statistics (AISTATS), 2020.
- P Pruthi, J González, **X. Lu**, M Fiterau, Structure mapping for transferability of causal models, in Inductive Biases, Invariances and Generalization in Reinforcement Learning Workshop in ICML, 2020.
- B Balaji, P Christodoulou, **X. Lu**, B Jeon and J Bell-Masterson, FactoredRL: Leveraging factored graphs for deep reinforcement learning, in NeurIPS Workshop on Deep Reinforcement Learning, 2020.
- **X. Lu**, Modelling, Inference and Optimization in Probabilistic Machine Learning, PhD thesis, 2019.
- **X. Lu**, Y. W. Teh, On Exploration, Exploitation and Learning in Adaptive Importance Sampling, arXiv preprint arXiv:1810.13296, 2018.
- T. Rainforth, Y. Zhou, **X. Lu**, Y. W. Teh, F. Wood, H. Yang and J. W. Van de Meent, Inference trees: Adaptive inference with exploration, Advances in Approximate Bayesian Inference, NIPS 2017 Workshop.
- **Lu, X.**, J, González, Z. Dai, N. Lawrence, Structured Variationally Auto-encoded Optimization, in International Conference on Machine Learning (ICML), 2018.
- **X. Lu**, V. Perrone, L. Hasenclever, Y. W. Teh, S. J. Vollmer, Relativistic Monte Carlo, in Artificial Intelligence and Statistics (AISTATS), 2017.
- **X. Lu**, J, González, Z. Dai, N. Lawrence, Probabilistic Optimization with Latent Search for Automatic Model Selection, 7th NIPS Bayesian Optimization workshop in Science and Engineering, 2017.
- H. Kim, **X. Lu**, S. Flaxman, Y. W. Teh, Tucker Gaussian Process for Regression and Collaborative Filtering, in Women in Machine Learning Workshop (WiMl), 2016.

Selected Internship and Teaching Experience

07/2018 - 09/2018, Microsoft Research, Research Intern, UK

- Worked on a Reinforcement Learning research project using imitation learning with latent variable models, supervised by Jan Stuehmer and Katja Hofmann. We use a generative model to capture different emergent play styles in video games, enabling the imitation of a diverse range of distinct behaviours in Minecraft. More details can be found here.

09/2017 - 11/2017, JP Morgan Chase, Quantitative Research Intern, UK

- Worked on model review on CDS risk that are not captured by VaR (Value at Risk), including restructuring and quanto effect. Frequently delivered high quality model review reports.
- Tested model assumption and data validation with statistical back testing.

06/2017-08/2017, Amazon, Applied Scientist Intern, UK

- Delivered research project in Bayesian Optimization when the input space is non-Euclidean, with an application in automated model selection and natural scene understanding. Successfully published the paper at ICML, 2018.
- Implemented the VAE (Variational Autoencoder) module in a deep learning framework (MxNet) and contributed to the open-sourced MxNet repository.

10/2016-06/2017, St Hilda's College, University of Oxford, Non-stipendiary Lecturer, UK

- Teaching Probability & Statistics course to undergraduate students, marking scripts and assisting with the undergraduate admission process.
- Improved communication and presentation skills, gained leadership and academic admission experience.

Skills

- Proficient with Python and machine learning libraries such as TensorFlow, etc..
- Example open source code on explainable models available here.
- Proficient with AWS tools and pipelines such as EC2, Sagemaker notebooks, state machines, S3 storage, lambda function etc.
- Experienced with big data management and tools such as SQL and ETL.

Scholarships and Awards

- Clarendon titular scholarship, 2014 - 2019.
- PAG Oxford Scholarship, 2014 - 2019.
- Royal Statistical Society Prize, 2014.
- Gibbs Prize, University of Oxford, 2014.
- Department of Statistics Prize, 2013.
- Department of Statistics Prize, University of Oxford, 2013.
- Top ten finalist for the Mathematics, Economics and Finance Undergraduate Of The Year Award, 2013 by TARGETjobs.