Xiaoyu Lu

luxiaoyu644@gmail.com
Personal Website: https://xiaoyulu2022.github.io/
LinkedIn: https://www.linkedin.com/in/xiaoyu-l-a449616b/
Google Scholar: https://scholar.google.com/citations?hl=en&user=ELqAe3MAAAAJ

About Me

I have worked as a machine learning scientist at Amazon for 4 years, where I have developed multiple innovative initiatives using machine learning models. Recently, I have been working on explainable AI to emulate complicated systems. I enjoy owning and driving projects forward, from initial scoping to implementation to make a real life impact. I have a track record of successful delivery of ML/Statistical solutions to business problems. I gained leadership through mentoring other scientists and leading our business intelligence team. I am proficient with Python, AWS tools including cloud computing, SageMaker, State Machine etc.. I am also proficient with data analytics tools such as data pipelines and SQL. Prior to Amazon, I completed my PhD in probabilistic machine learning 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 in Mathematics and Statistics at University of Oxford with the MMath degree, during which I topped the department in both the bachelor (3rd year) and the master year (4th year).

Work Experience

04/2019 - Present, Machine Learning Scientist, Amazon, UK

- I have been working on building explainable emulators for complex systems, providing interpretable attribution on key business metrics. My work on the explainable model has been published at the International Conference on Machine Learning (ICML) 2022.
- I have innovated and developed multiple initiatives in Amazon. For example, I developed a tool to explain evolution of key business metrics in Amazon's supply chain systems. I have successfully deployed the tool in production since 2022.
- I acquired leadership through supervising and mentoring applied scientist interns and data analysts, continuously developing the business intelligence capabilities of my team.
- I have experience working with big data and utilising cloud computing. I write production-level Python code and I am proficient at AWS tools, data analytics, SQL and ETL.
- I lead and drive projects from scoping to launching, building technologies that are best suited to the business problems under ambiguous environments. I developed my stakeholder management through regular communications of project progress, proactively pushing projects forward with customer obsession.

Education

10/2014 - 06/2019, PhD in Probabilistic Machine Learning, University of Oxford, UK

- I obtained my PhD degree working in the field of machine learning with Prof. Yee Whye Teh. I have research experience in probabilistic generative models, Bayesian inference, Gaussian Process, MCMC, causal inference and reinforcement learning, with publications at prominent AI/ML 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.
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, arXiv preprint arXiv:1806.09550, 2018.
- **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.
- 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

- I was working 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 an unsupervised manner, 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 back testing.

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

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

07/2015-10/2015, Google, Technical Business Intern, Ireland

- Created competitive analysis and benchmarking study for account hijacking, recommending strategy adjustments based on findings.
- Analysed hijacking trends within a specific set of products and developed an action plan based on trends and patterns.
- Analysed preventable abuse related issues which impact users, and identify core and common prevention focus areas across Product Quality Operation.
- Partnered with engineering teams to improve hijacking prevention, detection and recovery systems.

06/2014-08/2014, Credit Suisse, Quantitative Strategies Summer Analyst, UK

- Built pricing models for Calendar Spread Options using Excel and VBA.
- Performed model calibration and validation, as well as hedging simulation for historical data.
- Delivered excellent results and received exceptional feedback from managers and colleagues.

Skills

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

Scholarships and Awards

- Clarendon titular scholarship.
- PAG Oxford Scholarship.
- 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.