# Xiang (Shawn) Liu

**Machine Learning Engineer**

shawnliuxiang@gmail.com | 778-881-7753 https://www.linkedin.com/in/shawnliuxiang/


## TECHNICAL SKILLS

- Programming languages: Java, Python, Scala, C/C++, Julia, MATLAB, R, TypeScript.
- Framework and platforms: PyTorch, Hugging Face, Keras, Scikit-learn, NLTK, Docker, Flask.
- AWS: EC2, Lambda, Batch, StepFunction, SageMaker, S3, DynamoDB, SQS, EventBridge, CloudFormation, CloudWatch.

## EXPERIENCE

**Machine Learning Engineer**, **Amazon Search Science and AI**, Vancouver Aug 2022 – Present

- Served as the lead MLE in a 20-member team of software engineers and mentored 2 engineers in transitioning to MLE roles;
- Led the "SuperFresh" program, enhancing the Amazon LLM shopping experience with the latest shopping-related information;
- Led Amazon's internal shopping LLM pre-training data processing pipeline, leveraging on-demand global AWS computing infrastructure to execute resource-intensive tasks while managing costs effectively;
- Boosted LLM-based data fluency improvement workflow by 175x, reducing a single data release cost from $10M to $0.1M;
- Fined-tuned multiple LLMs, including pre-training shopping data fluency improver, internal data anonymization processor, customer review summarizer, and reinforcement learning from human feedback (RLHF) preference data generator;
- Developed a dietary preference prediction pipeline leveraging grocery shopping history and achieved precision of 39%;
- Prototyped a smart basket widget in Indian Amazon Fresh to pre-fill the shopping cart with frequently ordered items;
- Partnered with Deloitte to elucidate team's machine learning projects, resulting in securing a $1.5M Canadian SR&ED tax credit.

**Software Development Engineer**, **AWS Identity**, Vancouver Aug 2020 – Jul 2022

- Led 3 engineers to launch a consistency verifier, driving designs, developments, compliance reviews, and operational work;
- Improved the efficiency of the data consistency verifier by skipping 95% of unnecessary traffic, saving $665K annually;
- Designed and developed a production identity database backfill pipeline, enabling secure reuse without manual intervention;
- Extended the search and query functions of the NoSQL-based identity service by adding Elasticsearch support;
- Simplified the test and release process with Infrastructure as Code and reduced service build in 28 AWS regions to only 1 day.

**Graduate Research Assistant**, **University of British Columbia**, Vancouver Sep 2018 – Apr 2020

- Led a large-scale project (40TB) analyzing the human brain network using high-resolution functional MRI;
- Accelerated neuroimaging analyses by 30x using cloud-based parallel computing on an advanced research computing system;
- Developed a temporal convolutional network to measure depression severity using only facial expression and spoken language.

**Software Engineer Intern**, **AltumView (an AI & smart home device startup)**, Burnaby May 2018 – Aug 2018

- Collaborated on building a smart mirror prototype, serving as a personal skin health tracker and beauty assistant in C++;
- Enhanced the CNN (Inception v3) based skin cancer classification model to achieve an accuracy of 84%;
- Created algorithms for hair removal, pole detection, skin texture detection, and oiliness measurement using OpenCV.

## PUBLICATIONS AND AWARDS

- Ge, R., Liu, X., Long, D., Frangou, S., & Vila-Rodriguez, F. (2021). Sex effects on cortical morphological networks in healthy young adults. NeuroImage, 233, 117945. https://doi.org/https://doi.org/10.1016/j.neuroimage.2021.117945
- Liu, X. (2020). Machine learning based prediction of repetitive transcranial magnetic stimulation treatment outcome in patients with treatment-resistant depression. https://open.library.ubc.ca/collections/24/items/1.0392572
- International Student Tuition Award, University of British Columbia
- Faculty of Medicine Graduate Award, University of British Columbia

## EDUCATION

**Georgia Institute of Technology**, Online

- Master of Science, Computer Science; GPA: 4.0/4.0; Specialization: Machine Learning. Dec 2023

**University of British Columbia,** Vancouver

- Master of Science, Experimental Medicine; GPA: 90%; Research Area: Medical Machine Learning; May 2020
- Bachelor of Science, Computer Science and Statistics. May 2018