

MICHAEL LIU

+1(240) 620-5618 ◇ New York, NY

michaelyliu6@gmail.com ◇ [linkedin.com/michael-liu](https://www.linkedin.com/in/michael-liu) ◇ michaelyliu6.github.io

SKILLS

Programming	Python, Java/Kotlin, C++, SQL, TypeScript, Go
Cloud Services	AWS (EC2, S3, DynamoDB, CloudFormation, SQS, SNS, IAM, Lambda)
Data Processing	Large-scale distributed systems, high-throughput data pipelines, CI/CD pipelines
AI/ML	Transformer, RL, Evaluation, Agents, Multimodal, PyTorch, Einops, TensorFlow, WandB

EXPERIENCE

Software Development Engineer II July 2025 - Present
Amazon Web Services (AWS) - AWS Security *New York, NY*

- Building a new team to centralize all physical security device information and management within Amazon

Software Development Engineer II Aug 2022 - Oct 2024
Amazon Web Services (AWS) - EC2 Core Platform *Herndon, VA*

- Oversaw global AWS data transfer metering—processing 140 PB of data across 12 million batches daily and generating \$11 million in daily revenue
- Ensured 24/7 system reliability through on-call rotations that resolved critical outages, unblocked CI/CD pipelines, and managed customer onboardings and inquiries
- Led a team of 4 engineers in implementing zero-touch automation for new AWS region builds—coordinating tooling adoption and cross-team efforts to accelerate data center build outs
- Designed and executed a S3 storage optimization project that cut costs by 90% on the team's largest S3 bucket, saving ~600PB/month and boosting operational efficiency

PROJECTS

Transformer Implementations and Mechanistic Interpretability Analyses. Trained GPT-2 transformer from scratch with various sampling techniques integrated with KV cache. Used TransformerLens to reverse engineer model behaviors (attention patterns, induction circuits). ([View in GitHub](#))

Reinforcement Learning Experiments. Implemented classical and deep reinforcement learning algorithms (SARSA, Q-learning, DQN, PPO, RLHF) in PyTorch and compared their performances in different gym environments. ([View in GitHub](#))

LLM Evaluations and Agent Experiments. Designed threat models and ran evaluations to test for model behaviors. Built LLM Agent scaffolding to play Wikipedia Racing Game using different elicitation methods (ReAct, Reflexion, etc). ([View in GitHub](#))

Diffusion and Multimodal Models Implementation. Trained diffusion models from scratch with various sampling methods (DDPM, DDIM) and implemented inference for multimodal architectures (CLIP). ([View in GitHub](#))

CERTIFICATIONS

- Machine Learning Engineering for Production (MLOps) Specialization - DeepLearning.AI (Feb '24)
- LLM101/102x: Foundation Models/Application Through Production - Databricks (Dec '23)
- TensorFlow Developer Certificate - Google (Oct '22)
- Deep Learning, Natural Language Processing Specialization - DeepLearning.AI (Jun '22, Jun '23)

EDUCATION

B.S. in Computer Science, University of Maryland, College Park
Machine Learning Specialization

Aug 2018 - May 2022