

Aryan Sarswat

[Portfolio](#) | [Email](#) | [LinkedIn](#) | [Github](#)

EXPERIENCE

Machine Learning Scientist II

Expedia Group

Jul 2025 – Present

Seattle, WA

- **LLM Evaluation Framework:** Architected multi-stage extendable evaluation pipeline for tool call classification, argument validity, hallucination detection, and instruction following; built system to reverse-engineer user prompts from production Splunk logs (MCP tool arguments) to generate **1K+ sample synthetic dataset** for regression testing without access to raw OpenAI queries; designed analysis dashboard adopted as **company-wide AI quality framework** across multiple teams
- **PromptLab:** Developed internal agentic prompt engineering platform enabling technical & non-technical users to iterate on prompts with auto-generated test cases, evaluation-driven improvements, and multi-LLM provider support
- **ChatGPT MCP Server Integration:** Led development during OpenAI partnership; drove **33+ PRs** across tool development, schema design, and compliance; built hotel, flight, and property info tools with prompt optimization and guardrails. Peaked at **15K users** within 2-week sprint
- **Flight Delay Prediction:** Engineered end-to-end system processing **78M+ rows**; trained production-ready CatBoost model achieving **89.9% accuracy** impacting **38M+ passengers**; operationalized FastAPI real-time inference pipeline; delivered revenue impact modeling analyzing **1.4M+ flights**
- **Hackathon Semi-Finals:** Led team to semi-finals; developed memory system using AI agents for personalized customer interactions; currently transitioning to production

Machine Learning Science Intern

Expedia Group

May 2024 – Jul 2024

Seattle, WA

- Developed **Image2Text pipeline** using VLMs (PaliGemma, BLIP-2) with quantized Llama3; deployed to production for image ranking, generating **\$5.4M GP** and **\$1.85M travel ad spend uplift**
- Built Image2Video pipeline (Dynamicroft, Swin2SR, FLAVR) and Image Relevancy pipeline (CLIP, BLIP) with custom search for destination imagery matching

Graduate Teaching Assistant – Deep Learning

Georgia Institute of Technology

Jan 2024 – May. 2025

Atlanta, GA

- Created assignments for building transformers from scratch; held weekly office hours; mentored and graded 5 course projects per semester

Research Assistant

Agency for Science, Technology and Research (A*STAR)

Aug 2022 – July 2023

Singapore

- Developed state-of-the-art deepfake detection models achieving **10% improvement** over existing methods on novel datasets; formulated scalable video processing pipeline handling datasets up to **1TB**

Teaching Assistant – Machine Learning

National University of Singapore

Aug 2022 – Dec. 2022

Singapore

- Taught weekly classes of 15 students on ML topics; mentored 3 project groups; involved in setting and grading assignments, exams, and final projects

EDUCATION

Georgia Institute of Technology

Master of Science in Computer Science; GPA: 4.0/4.0

Atlanta, GA

Aug 2023 – May 2025

- Concentration: Artificial Intelligence, Perception, and Natural Systems
- Coursework: Deep Learning, NLP, Reinforcement Learning, Computer Vision, Efficient ML

National University of Singapore

Bachelor of Computing with Distinction; Specialization in Artificial Intelligence

Singapore

Aug 2019 – May 2023

- University Scholar Programme (USP): Multidisciplinary honors program for top 1% of students

SELECTED PROJECTS

Cobalt: Scalable Teleoperation for Robot Learning Masters Thesis	Jan 2024 – May 2025
<ul style="list-style-type: none">Developed teleoperation platform to democratize robot learning at scale in simulation and real worldLeveraged vectorized environments with load-balanced infrastructure supporting concurrent multi-user teleoperation on single GPU	
Efficient Backdoor Unlearning in Language Models	Jan 2025 – May 2025
<ul style="list-style-type: none">Explored early-stage detection and removal of backdoors in LMs without retraining using SPECTRE's robust covariance estimation adapted to language modeling contextPerformed gradient ascent on detected poisoned samples; demonstrated early intervention reduces ASR while maintaining model accuracy with significantly lower computational overhead	
Efficient-Selection-via-Pruning (ESVP)	Jan 2024 – May 2025
<ul style="list-style-type: none">Novel data selection approach using model pruning and quantization for active learning; achieved 1.2-2x speed-up with structured pruning, 1.35x with quantization, learning with 50% less data at only 4% F1 loss	
Zero-Shot Adaptation for Guided Robot Policies	Aug 2024 - Dec 2024
<ul style="list-style-type: none">Introduced framework combining diffusion-based trajectory planners with LLM-generated loss functions for zero-shot policy adaptation; achieved real-time adaptation to spatial and speed constraints in MetaWorld	
Historical Document Restoration with GANs	Jan 2025 – May 2025
<ul style="list-style-type: none">Trained Pix2Pix GAN to reconstruct damaged documents; fine-tuned TrOCR for baselines; demonstrated GAN-based restoration significantly improves OCR accuracy on simulated degradations	
Deep RL in 3D Connect 4	Jan 2024 – May 2024
<ul style="list-style-type: none">Investigated DQN, Actor-Critic, and PPO via self-play; achieved 10-20% success rate against Minimax depth-2, consistently outperforming random and shallow Minimax agents	

HONORS & AWARDS

Hackathon Semi-Finalist – Expedia Group	2025
Hackathon Microsoft, GenAI, and Assurant Challenge Winner – Georgia Institute of Technology	2024
Certificate of Distinction in AI – National University of Singapore	2022
USP Senior Honour Roll – National University of Singapore	2021