Shinhaeng Lee

Atlanta, GA | shinhaenglee.com | slee3487@gatech.edu | 470-676-4493

Education

Georgia Institute of Technology | Atlanta, GA

B.S Computer Science, GPA 3.7 / 4.0 — *Dec 2025*

M.S Computer Science, Specialization in Machine Learning — *Expected Dec 2026*

Experience

UN Peacekeeper - United Nations Interim Force in Lebanon (UNIFIL) *June 2023 - January 2024*

- Maintained, repaired, and constructed websites and servers for internal communication, and managed computer devices as part of the Signal Company, Republic of Korea Battalion.
- Engaged in civilian operations and peacekeeping initiatives to support the local community and worked with multinational forces to enhance regional stability and security.

Undergraduate Researcher - The Teachable AI Lab (Georgia Institute of Technology) *August 2025 - present*

- Invented a **pipeline integrating LLMs with Decomposed Inductive Procedure Learning (DIPL)** — where LLMs extract observations from GSM8K problems, and DIPL (a symbolic program-induction / rule-learning framework) enables agents to induce production rules (skills) from worked examples with explicit *how/where/when* mechanisms.
- **simulating human-like learning** to attain greater robustness and efficiency over standard LLMs in mathematical word problem solving by leveraging symbolic reasoning with LLM-based observations.

Undergraduate Researcher - Sarker Lab (Emory University) *May 2025 - present*

- Architected **a modular NLP framework** (regex + lexicons + medication-tuned PLMs + LLMs) to extract and classify polysubstance-use from Reddit.
- Modeled **multi-order Markov chains** to quantify month-to-month transition probabilities between substances.
- Leveraged **Apriori** and **PrefixSpan** to mine frequent and sequential substance use patterns while preserving co-occurrence and temporal order

VIP Research Team (BTAP) - Georgia Institute of Technology *January 2025 - December 2025*

- Developed an **iOS application providing Augmentative and Alternative Communication (AAC) solutions** for individuals with traumatic brain injuries (TBIs), contributing to UI/UX design and core feature implementation, including Visual Scene Display, text-to-speech, and AI-driven smart suggestions.
- Collaborated in a multidisciplinary team to research and implement accessibility-focused technology, leveraging Swift and SwiftUI to enhance communication tools for individuals with aphasia, while addressing challenges such as data persistence and usability improvements.

Personal Projects

Real-Time YOLOv3 Object Detection Model for Gastrointestinal Endoscopy *August 2024*

- Built a **real-time YOLOv3 model** with PyTorch and OpenCV for gastrointestinal polyp detection, fine-tuned on the **Kvasir dataset** (mAP@0.5 = 0.74).
- Optimized model using **L1 and Taylor Expansion pruning**, reducing parameters by **92% with only 11% accuracy loss**, and achieving **20% greater reduction** than L1 alone.
- Applied **K-Means clustering** for anchor box optimization, accelerating convergence by **40%**.
- Implemented a **custom greedy pruning strategy**, improving real-time inference without accuracy loss.
- Achieved **19 FPS at 54ms latency on CPU**, enabling efficient deployment on lower-performance hardware.

Fine-Tuning CoT: GRPO vs. PPO vs. Few-Shot Prompting *March 2025*

- Implemented **GRPO, PPO, and CoT few-shot prompting from scratch** to compare fine-tuning approaches for Chain-of-Thought reasoning in **GPT-Neo** using the **GSM8K dataset.**
- Trained a **Reward Model** by fine-tuning **DistilRoBERTa-base** on a comparison dataset to provide structured feedback for reinforcement learning and analyzed performance to determine the most effective fine-tuning strategy.

CycleGAN for Face-to-Portrait Image Generation *February 2024*

- Developed a custom **CycleGAN** model for **unsupervised transformation** of human faces into portrait images using **PyTorch**, incorporating **UNet-based encoders and decoders within PatchGAN** for enhanced architecture.
- Improved model performance by adding **Local Self-Attention** to focus on critical regions and utilizing **buffering of past generated images** to reduce training oscillation, ensuring stable and reliable convergence.

Autonomous Driving Simulation with Genetic Algorithm *December 2023*

- Simulated real-time road environments, vehicles, and sensors via a web-based interactive interface, enabling autonomous navigation using a **feedforward neural network** built with **JavaScript, HTML, and CSS.**
- Enhanced driving performance through **Genetic Algorithm,** applying **Elitism, Roulette Wheel Selection,** and**, Two-Point Crossover** for continuous performance improvement.

Skills

Languages: Python, R, java, C/C++, C#, SQL, Javascript, HTML/CSS

Frameworks and Libraries: PyTorch, Tensorflow/Keras, Scikit-learn, Matlab, Numpy, Pandas, Matplotlib, SpaCy, OpenCV, Next.js

Technologies: Git, Linux, Docker, LLVM, CUDA, AWS, NLTK, Tableau, Power BI