

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

M.S Computer Science, Specialization in Machine Learning, GPA 4.0 / 4.0

Dec 2025

Expected Dec 2026

Research Experience

Undergraduate Researcher - The Teachable AI Lab (Georgia Institute of Technology)

August 2025 - present

- Designed a neuro-symbolic tutoring pipeline that couples **LLM-based semantic parsing** of **GSM8K** word problems with **Decomposed Inductive Procedure Learning (DIPL)** in **TutorGym**, enabling agents to **induce interpretable production-rule policies** from worked-example demonstrations via explicit **how/where/when** mechanisms.
- Improved robustness vs. pure LLM reasoning by inducing modular algebra skills (e.g., substitution, isolation, arithmetic) via explicit **how/where/when** rule mechanisms.

Undergraduate Researcher - Sarker Lab (Emory University)

May 2025 - present

- Developed a modular NLP system (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.
- Mined frequent and sequential patterns using **Apriori** and **PrefixSpan** while preserving co-occurrence and temporal order.

Experience

UN Peacekeeper - United Nations Interim Force in Lebanon (UNIFIL)

June 2023 - January 2024

- Maintained and deployed internal websites/servers supporting battalion communications; administered computing devices for operational readiness.
- Supported civil-military operations and peacekeeping initiatives with multinational units.

VIP Research Team (BTAP) - Georgia Institute of Technology

January 2025 - December 2025

- Built an iOS **AAC** app for users with TBI/aphasia; implemented features including **Visual Scene Display**, **text-to-speech**, and **AI-based smart suggestions**.
- Worked in a multidisciplinary team to improve accessibility and usability; contributed to **Swift/SwiftUI** development and data persistence.

Personal Projects

Real-Time YOLOv3 Object Detection Model for Gastrointestinal Endoscopy | PyTorch, OpenCV

August 2024

- Fine-tuned YOLOv3 on the **Kvasir** dataset; achieved **mAP@0.5 = 0.74**.
- Reduced parameters **92%** via L1 + Taylor pruning with **~11%** accuracy loss; exceeded L1-only reduction by **20%**.
- Applied **K-Means clustering** for anchor box optimization, accelerating convergence by **40%**.
- Optimized anchor boxes with k-means (40% faster convergence) and reached **19 FPS / 54 ms latency on CPU**.

Fine-Tuning CoT: GRPO vs. PPO vs. Few-Shot Prompting | GPT-Neo, GSM8K

March 2025

- Implemented **GRPO**, **PPO**, and few-shot CoT prompting from scratch; evaluated reasoning performance on GSM8K.
- Trained a reward model by fine-tuning **DistilRoBERTa-base** on preference comparisons to provide structured RL feedback.

CycleGAN for Face-to-Portrait Image Generation | PyTorch

February 2024

- Built a CycleGAN with UNet-style encoders/decoders + PatchGAN; added local self-attention and image buffering to stabilize training.

Autonomous Driving Simulation with Genetic Algorithm | JavaScript, HTML/CSS

December 2023

- Built a browser-based driving simulator; optimized a neural-network controller using elitism, roulette-wheel selection, and two-point crossover.

Skills

AI/ML: Hugging Face (Transformers, Datasets), BERT/BioBERT, PEFT/LoRA, FAISS, Deep Learning, Generative AI, NLP, Computer Vision; PyTorch, TensorFlow/Keras, scikit-learn

Data & Analytics: Python, R, SQL; NumPy, Pandas; Statistical Modeling, Feature Engineering, Predictive Analytics; Matplotlib/Seaborn; Tableau, Power BI

Deployment & Tools: AWS, Docker, FastAPI, Streamlit, Jupyter/Colab, MLOps, Model Deployment; Git, Linux, CUDA, LLVM

Languages/Other: Java, C/C++, C#, JavaScript, HTML/CSS; MATLAB; spaCy, NLTK, OpenCV; Next.js