

Eunbi Yoon

Email: eunbiyoon@umass.edu | Address: Amherst, MA, United States | Phone: +1 413-399-7241
LinkedIn: [linkedin.com/in/eunbiyoon0717](https://www.linkedin.com/in/eunbiyoon0717) | GitHub: github.com/EunbiYoon | Portfolio: www.eunbiyoon.com

EDUCATION

University of Massachusetts Amherst | Master's Degree in Computer Science, AI/ML | Sep 2024 – Dec 2026

- **GPA:** 3.93, **Coursework:** AI Alignment, Neural Networks, Machine Learning, Artificial Intelligence, Reinforcement Learning

Kyung Hee University | Bachelor's Degree in Electronics Engineering, Computer Engineering Core | Mar 2016 – Feb 2020

PAPERS / PROJECTS

TRACE: Separating Action and Reasoning Teachers for Strategic SLM Training | Co-first Author, with Amazon Science, Accepted at NeurIPS 2026 SLM Agent Workshop

- Fine-tuned Qwen2.5-3B by combining generated preference datasets, distillation from separate action and reasoning teachers, and DPO with LoRA to improve strategic reasoning across 12 multi-step, multi-agent tasks.
- Evaluated the preference optimized 3B model in in-domain, held-out, and strategic-opponent settings, raising tit-for-tat reward from 1.50 to 14.06, a 9× gain surpassing frontier Haiku 4.5 and much larger Gemma 3-27B and Llama 3.3-70B.

GAMEBIT: Grounded Adaptive Modeling of Explicit Beliefs for Interactive Training | First Author, UMass Resource-Bounded Reasoning Lab (Advisor: Prof. Shlomo Zilberstein), Target: ICML 2027 Conference

- Post-trained Qwen3-4B agents with GRPO through self-play against hidden-identity opponents, designing proper scoring rule rewards for calibrated opponent beliefs and coherence rewards so agents act on what they reason.
- Debugged distributed RL training on multi-node GPUs with ROLL, pairing vLLM rollouts with Megatron-Core training via Ray, tracing anomalous runs to root causes through think-belief-plan-action traces and memory profiling against OOM.

Who Gets the Credit? Reward Design and Good Teammates in Cooperative RL

- Applied RLAI-style reward learning, training Bradley-Terry reward models on LLM-as-a-judge preferences over trajectories, post-training collaborating agents with PPO and GRPO against learned, team, and individual rewards.
- Audited LLM-as-a-judge and learned-reward integrity against counterfactual Shapley contribution, measuring how often judges credit busy-looking agents and whether agents trained on learned rewards learn fake help or credit theft.

SKILLS

Post-Training: SFT, RLHF, DPO, GRPO, PPO, Reward Modeling, LoRA, PyTorch, JAX, Hugging Face (Transformers, TRL, PEFT)

Evals & Inference: Agentic Evals, LLM-as-a-Judge, Reward Hacking Detection, Synthetic Data, vLLM, Quantization

Distributed Training: FSDP, DeepSpeed, Megatron-Core, Ray, Tensor/Pipeline Parallelism, Multi-Node GPU Clusters, CUDA

Data & Infrastructure: Python, SQL, Spark, Airflow, Docker, Kubernetes, AWS (S3, EKS, SageMaker, Bedrock), Git

PROFESSIONAL EXPERIENCE

Backend Engineer 2 | PTKOREA (Samsung's Subsidiary) | Data Engineering Team | Feb 2025 – Oct 2025

- Developed production agentic RAG assistant over LangGraph, LlamaIndex, and Chroma vector database, designing labeled eval sets with grounding and hallucination checks and low confidence abstention that cut ungrounded answers by 30%.
- Built distributed pipelines processing billion-scale Samsung.com traffic and conversion data from Adobe Analytics APIs, with Spark ETL on AWS Glue, Parquet to S3, and Airflow on EKS, powering customer behavior analytics.
- Designed customer facing monitoring dashboard for Airflow pipelines on Docker and Kubernetes, letting stakeholders track output data quality and pause or resume pipelines in real time, cutting incident response time by 50%.

Software Engineer 2 | LG Electronics North America | R&D Team | Jan 2020 – Jun 2024

- Developed scanned image defect detection models with TensorFlow and OpenCV in production, monitoring data drift and retraining on misclassified samples, while upholding code quality through code reviews and ML knowledge sharing.
- Automated support and approval workflows, developing customer self-repair platform on ECS via GitLab and Docker CI/CD that cut service calls by 30%, and employee time-off and security platform synced with HR, cutting processing time by 80%.

CERTIFICATIONS

- AWS Certified Machine Learning Engineer Associate (MLA-C01) — Amazon Web Services, Sep 2026
- Deep Learning with TensorFlow – Advanced (ML0120EN) — IBM, Sep 2026
- AI Agent Fundamentals & Generative AI Fundamentals — Databricks, Jun 2026