Subin Park

Seoul, Republic of Korea | enlightkorean@gmail.com | SOL1archive.github.io
linkedin.com/in/subin-park-605560278 | github.com/SOL1archive

Education

Yonsei University, Seoul, BS in Computer Science — Mar 2022 – Present

- **GPA:** 3.93/4.3
- **Major:** Computer Science
- **Main Coursework:** Software Engineering, Computer Network, Computer System, Intro. to Computing Research, Deep Neural Networks, Mathematics for Deep Learning, Engineering Mathematics I & II, Linear Algebra, Probability & Statistics, Text Mining, Intro. to Cognitive Neuroscience

Research Interests

My ultimate research goal is to delve into the fundamental question: "Can machines think?" More specifically, I aim to comprehend the very essence of human intelligence and implement a generalized framework of intelligence that is grounded in mathematics, linguistics, and cognitive neuroscience.

- **Natural Language Processing:** LLM Alignment, Transformer-Alternative Architectures (e.g., Mamba, Titans), LLM Reasoning
- **Reinforcement Learning:** Offline RL, Sequential Decision Making Problem, Data-Efficient RL
- **Foundation Models:** NLP, RL, Causal Foundation Model

Experience

Undergraduate Researcher, Decision Making Intelligence & Learning Lab, Yonsei University — Apr 2025 – Present

- **Advised by:** Prof. Jongmin Lee
- Leading research on Offline Direct Multi-Preference LLM Alignment via RLHF
- Participating in research on the RL foundation model for world modeling
- Applied and implemented various RL algorithms (e.g., SAC, PPO)

Military Intelligence Specialist (Sergeant), Republic of Korea Army — Jul 2023 – Jan 2025

- Gathered, analyzed tactical intelligence to support unit operations (Specialized in HUMINT, OSINT)
- Led a squad consisting of Intelligence Specialists and Operations Clerks (More than 6 months)
- Completed CBRN threat analysis and mission planning training

Awards & Scholarships

Sangah IT Talent Scholarship — Jul 2025 – Jun 2027 (Expected)

- **Awarded by:** Sangah Scholarship Foundation
- Awarded a scholarship of 5,000,000 KRW (≈ 3,500 USD) per semesters for up to 3 years

2025-1 Veritas Scholarship — Feb 2025

- **Awarded by:** Yonsei University
- Awarded to students with outstanding GPA

2023-1 Academic Excellence Award — Aug 2023

- **Awarded by:** Yonsei University
- Awarded to the top 10% of students by GPA
- **Courses:** Object-Oriented Programming, Discrete Mathematics, Text Data Processing, Introduction to Economics, Introduction to Brain Engineering, Introduction to Data Science, Introduction to Cognitive Neuroscience

Yonsei Humanities and Arts Vision Excellence Team Award Jun 2023

- **Awarded by:** Yonsei University
- Awarded to leaders of interdisciplinary teams bridging Humanities and Technology

Projects

PyChain Lab Jan 2025 – Dec 2025

- Investigated Chain-of-Thought (CoT) reasoning enhancement and Diffusion LLM architectures.
- Conducted research on reasoning capabilities of Large Language Models.
- Tools Used: Python, PyTorch, Transformers, Diffusion Models
- **Link:** github.com/SOL1archive/pychain-lab

Multi-Agent RL-based Stock Trader Aug 2025 – Sept 2025

- Led the development of a multi-agent RL-based stock trader. Actively searched for related research and setting project objective for novelty and high performance.
- Achieved 16% earning rate in the test time.
- Awarded 2nd Position in the competition at Yonsei Univ. Data Science Lab.
- Tools Used: PyTorch, Gymnasium, TradeRL
- **Link:** github.com/DataScience-Lab-Yonsei/25-2_DSL_Modeling_RL_Stock_Trader

Joseon Document Translator Apr 2025 – Jun 2025

- Led development an ancient-to-modern Korean translator by fine-tuning Qwen2.5-1.5B-Instruct; deployed the model on HuggingFace for accurate historical text translation
- Achieve SOTA performance on long context ancient document translation
- Tools Used: Python, Transformers, HuggingFace, Apache Arrow
- **Link:** github.com/SOL1archive/ancient-document-LLM

Learned Optimizer & Learning Rate Scheduler Mar 2025 – Apr 2025

- Developed a Transformer-based meta-learning learning rate scheduler that learns to adapt update rules automatically, significantly reducing manual hyperparameter tuning and accelerating convergence on MNIST benchmarks
- Tools Used: Meta-Learning, Reinforcement learning, Python, PyTorch, Transformer
- **Link:** github.com/DataScience-Lab-Yonsei/25- 1_DSL_Modeling_RL_Learned_Optimizer_LR_Scheduler

ClauseSummary Mar 2023 – Jul 2023

- Led development a clause-level summarization model using PPO-based RLHF to help investors understand complex financial contract terms; data crawling, preprocessing, model fine-tuning, and evaluation workflows
- Project Funded by the Institute of Higher Education Innovation Yonsei University (IHEI Workstation)
- Tools Used: Python, PyTorch, HuggingFace Transformers & TRL
- **Link:** github.com/SOL1archive/ClauseSummary

KoGrammar Apr 2023 – Jun 2023

- Led building a Korean grammar correction system atop KoBART with baseline, distilled, and tiny variants; implemented HuggingFace pipelines for inference and quantitative evaluation on the corpus from the National Institute of the Korean Language
- Tools Used: PyTorch, Transformers, Jupyter
- **Link:** github.com/SOL1archive/KoGrammar

Financial Data Noise Analysis Feb 2023 – Jun 2023

- Investigated and compared denoising methods (MA, EMA, bilateral filter, convolutional autoencoder) on financial time-series; performed noise distribution analysis via KL divergence, KDE, and violin plots
- Tools Used: Python, Jupyter, pandas, Matplotlib
- **Link:** github.com/SOL1archive/Fin-Data-Denoising

Activities & Participations

Yonsei Computer Club Member Aug 2022 – Jun 2023

- Project Lead: ClauseSummary
- Participated in Group Study: Data Analysis, SQL, Deep Learning

Yonsei AI Associate Research Member Jan 2025 – Dec 2025

- Participated in NLP paper review groups.
- Gave an open seminar for the NLP paper & RL theory presentation.
- Participated in PyChain Lab project on CoT reasoning.

Yonsei Data Science Lab Member Jan 2025 – Dec 2025

- Project Lead: Multi-Agent RL-based Stock Trader
- Participated in Exploratory Data Analysis for insurance data analysis and sports.

Open Source Contributions

EleutherAI Polyglot Project Member Apr 2023 – Jun 2023

- Participated as a NLP Researcher for Multilingual LLM
- Carried out Korean Chain-of-Thought Dataset Construction based on Crawling KSAT and other variational data
- Experienced Pre-train level massive data preprocessing and baseline LLM evaluation from reasoning benchmarks.
- Tools Used: Python, Jupyter, Spark, Hadoop Ecosystem, PyTorch, Transformers

Technologies & Skills

Programming Languages:

- **High:** C, Python
- **Middle:** Arduino, C++
- **Low:** Java, x86 Assembly

Technologies:

- **High**: PyTorch, HuggingFace (Transformers, TRL, Datasets), PyTorch Lightning, DeepSpeed, NumPy, Pandas
- **Middle**: Scikit-Learn, Gymnasium, Linux, Docker, Flask
- **Low**: Django, Spark, SQL

Other Skills: EDA (Exploratory Data Analysis), Statistical Analysis, Git, Bash
Languages: Korean(Native), English