Last updated in July 2024

Sangryul Kim

Seoul, Republic of Korea sangryul@kaist.ac.kr sangryul.github.io

Summary

I am a master's student at KAIST AI Graduate School in the xfact lab under Professor James Thorne, researching NLP, text retrieval, and deep learning. I enjoy creating tech products, with a focus on robust idea and developing MVPs. My work involves using large language models (LLMs) for search, retrieval, and SQL generation, and I am exploring efficient learning and data compression techniques. I have experience in full fine-tuning large models, building RAG tools with Korean models, and deploying models with Hugging Face, as well as skills in deep learning, data engineering, and web development.

Education

MS **KAIST AI** — Mar. 2023 to Feb. 2025(planned)
- GPA: 3.94/4.3
- Advisor: James Thorne / Explainable Factual Reasoning Lab(xfact)

BS **Chung-Ang University**, Major : Computer Science and Engineering / Minor : Applied Statistics — Mar. 2016 to Feb. 2023
- GPA: 3.98/4.5

Publications

ProbGate at EHRSQL 2024: Enhancing SQL Query Generation Accuracy through Probabilistic Threshold Filtering and Error Handling — May 2024
The 6th Clinical Natural Language Processing Workshop at NAACL 2024
Sangryul Kim, Donghee Han, Sehyun Kim

Re3val: Reinforced and Reranked Generative Retrieval — Mar. 2024
Findings of the Association for Computational Linguistics: EACL 2024
Euiyul Song, Sangryul Kim, Haeju Lee, Joonkee Kim, James Thorne

Experience

NAVER WEBTOON, Machine Learning Engineer, Intern — Seongnam-si, Republic of Korea, Jul. 2022 to Oct. 2022
- Developed an end-to-end program to block inappropriate ads by analyzing and visualizing data from ad images.
- Work Experience Field : Data Preprocessing, Statistical Machine Learning, Deploy and Manage the data and model

Cyber Operations Command of Republic of Korea, Discharged to the rank of sergeant — Seoul, Republic of Korea, Feb. 2019 to Sep. 2020
- Developed and operated military training program for cyber attack simulation with C++ and Windows API.
- Developed and operated malware and C2 server with C++, Python and Mimikatz.
- Developed a system for mobile security diagnosis using Vue.js, Node.js, ADB and MetaDefender.

ONYCOM Inc, Software Engineer — Seoul, Republic of Korea, Jan. 2018 to Jan. 2019
- Developed the pre-release version of IMQA MPM and Crash service as a Web programmer.

Projects

Gompada / github 2022

- Developed "Gompada," an AI grading tool for descriptive answers using SBert and FastText, focusing on keyword matching and contextual similarity.
- Designed and developed front-end (Vue.js) and back-end (FastAPI) services, optimizing ML model loading for faster response times.

Help Diana / github / demo 2021

- Developed a web service to bridge the information gap between doctors and patients using AI, Vue.js PWA, Spring, MySQL, AWS, and Docker.
- Led project planning, design, front-end development, AWS setup, data crawling, and API integration, including OCR and medical word tokenization.

AUTO TA / github 2021

- Developed a stock trading prediction project using FastAI, TensorFlow, PyTorch, FastAPI, and Docker.
- Led planning, model design, server development, data collection, model experimentation, and deploying models for stock prediction using CNN and LSTM, optimizing performance for real-time investment strategies.

Additional Experience

CIT Gangnam Apgujeong Coding Academy Instructor(Sep. 2021 to Mar. 2022): Worked as a coding academy instructor, prepares for overseas universities, and conducts Python, Data Science, Machine Learning classes and projects for adults

Certificated Boostcamp AI Tech 3th by Naver Connect Cooperation(Jan. 2022 to June 2022): Participated in NLP track, Got 1st place in the first team image classification competition on stages.ai, Experienced Text Retrieval, NER Downstream Task of NLP

Certificated Google Machine Learning Bootcamp by Google Developers(Aug. 2021 to Nov. 2021): Certificated Deep Learning Specialization(Coursera) by Andrew Ng, Participated in Kaggle Tabular Competition

Certificated SW Maestro 8th by Ministry of Science and ICT(Aug. 2017 to Dec. 2017): Developed “Mononity”, Perform mobile performance data extraction and analysis, Do extraction and diagnosis of Android smartphone performance information using ADB Shell

Awards

2nd Place: 2023 AI Graduate School Challenge hosted by Korea Ministry of Science and ICT

- Present "Explainable and Customizable Multihop Question Answering", Developed a RAG system with various custom features and integrated Korean language models.

Excellence Thesis Award: 2017 FALL KSII Conference

- "Analysis of Emotion Prediction Model by Genre Using Musical Features" / Sangryul Kim, Hyunki Lim. Dae-Won Kim