

Dongmin Choi

Seoul, Korea | chrisandjj1218@gmail.com | linkedin | homepage

Summary

I am a graduate student in Data Science at Seoul National University, researching **value alignment in LLMs** under Professor Yohan Jo. I believe aligning language models with specific value systems is a crucial first step toward building personalized LMs. I am also interested in improving reasoning capabilities in LLMs, particularly in mathematical domains.

Education

| | |
|---|---------------------|
| Seoul National University , M.S. in Data Science / Advisor: Prof. Yohan Jo | Mar 2024 – Present |
| Seoul National University , B.S. in Mathematics Education | Mar 2017 – Feb 2024 |

Experience

| | |
|--|----------------------|
| Undergraduate Intern , Institute of Convergence Technology at Korea Telecom, AI Service Planning Department – Korea | July 2023 – Aug 2023 |
| • Developed a voice lecture summary service utilizing KT E2E STT and GPT-3.5 technologies | |

Publications

| | |
|--|------|
| Value Portrait: Assessing Language Models' Values through Psychometrically and Ecologically Valid Items | 2025 |
| Jongwook Han*, Dongmin Choi *, Woojung Song*, Eun-Ju Lee, Yohan Jo | |
| [ACL] Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics, 2025, *Equal Contributions | |
| | |
| PVP: An Image Dataset for Personalized Visual Persuasion with Persuasiveness Ratings, Persuasion Strategies, and Viewer Characteristics | 2025 |
| Junseo Kim, Jongwook Han, Dongmin Choi , Jongwook Yoon, Eun-Ju Lee, Yohan Jo | |
| [ACL] Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics, 2025 | |

Preprints

| | |
|--|------|
| Established Psychometric vs. Ecologically Valid Questionnaires: Rethinking Psychological Assessments in Large Language Models | 2025 |
| Dongmin Choi , Woojung Song, Jongwook Han, Eun-Ju Lee, Yohan Jo | |
| Under Review at TACL | |
| | |
| Future Policy Aware Preference Learning for Mathematical Reasoning | 2025 |
| Minjae Oh, Yunho Choi, Dongmin Choi , Yohan Jo | |
| Under Review at ICLR 2026 | |

Projects

| | |
|--|----------|
| MCP-based Financial Q&A System | Nov 2024 |
| • Led idea development and overall project management | |
| • Implemented logic for selecting and invoking appropriate tools based on user queries using MCP and ReAct prompting | |
| • Built web demo using Streamlit | |
| • Tools Used: MCP, Streamlit, LangChain, Python | |
| | |
| Emotional Support Model with Targeted Data Augmentation | Oct 2024 |
| • Implemented pipeline for epoch-wise performance evaluation by topic and selective synthetic data augmentation | |

- Achieved comparable or superior performance using only 17% of synthetic data compared to full dataset usage

Voice Lecture Summary Service

July 2023 – Aug 2023

- Developed a service that summarizes lecture videos and generates slide notes after splitting content
- Implemented web frontend and integrated with KT E2E STT technology
- Utilized GPT prompt engineering to generate summaries, titles, correct STT errors, and perform post-processing
- Created training data for fine-tuning KT Mi:dm (LLM) for summarization tasks
- Tools Used: React, langchain

PDF Summary Service

July 2023

- Developed a web service that summarizes uploaded PDF documents after splitting them into appropriate chunks
- Deployed service using Vercel
- Tools Used: React, langchain, Vercel

KakaoTalk Open Chat Summary Service (YOUNG)

Feb 2023

- Developed a web service that summarizes KakaoTalk open chat conversations
- Researched and developed Dialogue Topic Segmentator (DTS) using unsupervised learning to address the lack of Korean datasets
- Built backend server using FastAPI
- Implemented model packing and server using BentoML
- Tools Used: FastAPI, BentoML