DongGeon Lee

M.S. student at POSTECH
Mail: dg.lee@postech.ac.kr
Web: https://donggeon.github.io

Research Interests

Natural language processing, Aligning Large Language Models (LLMs) to build trustworthy AI, Domain adaptation of LLMs

Educations

M.S. student in **Artificial Intelligence** — Feb 2024 - Present
Pohang University of Science and Technology (POSTECH) — *Pohang, South Korea*

- Advisor: Prof. Hwanjo Yu

B.S. in **Information and Communication Engineering** — Mar 2018 - Feb 2024
Inha University — *Incheon, South Korea*

Research Experiences

Graduate Research Assistant — Feb 2024 - Present
Data Intelligence Lab, POSTECH — *Pohang, South Korea*

- Advisor: Prof. Hwanjo Yu
- Research on knowledge conflicts of LLMs between external and internal knowledge.
- Research on continual domain-incremental learning in language models (LMs).

Undergraduate Research Assistant — Nov 2022 - Nov 2023
Data Intelligence Lab, Inha University — *Incheon, South Korea*

- Advisor: Prof. Wonik Choi
- Research on post-training of language models for domain adaptation.
- Research on keyphrase extraction from aviation incident reports via fine-tuning LMs.

Undergraduate Research Assistant — Jul 2021 - Jun 2023
Nursing Informatics Lab, Inha University — *Incheon, South Korea*

- Advisor: Prof. Insook Cho
- Research on fall event detection from electronic medical records via fine-tuning LMs.

Technical Skills

- **Programming Languages**: Python, Shell Script, (C++, C, JavaScript)
- **Frameworks and Libraries**: `PyTorch`, `transformers`, (`Keras`, `TensorFlow`)
- **Systems and Tools**: Git, Linux, LaTeX, (MySQL)

Conferences (International)

[1] Insook Cho, EunJu Lee, and **DongGeon Lee**. Effects of Language Differences on Inpatient Fall Detection Using Deep Learning. In the *Proceedings of the 19th World Congress on Medical and Health Informatics (MedInfo 2023)*, 2024.

[2] **DongGeon Lee**, EunJu Lee, and Insook Cho. Bridging the Reporting Gap of Inpatient Falls to Improve Safety Practices Using Deep-Learning-Based Language Models and Multisite Data. *AMIA 2023 Clinical Informatics Conference*, 2023.

Conferences (Domestic)

[1] **DongGeon Lee***, Ahjeong Park*, Hyeri Lee, Hyeonseo Nam, and Yunho Maeng. Question Types Matter: An Analysis of Question-Answering Performance in Retrieval-Augmented Generation Across Diverse Question Types. In the *Proceedings of the 36th Annual Conference on Human & Cognitive Language Technology (HCLT 2024)*, 2024.

[2] TaeYoon Kwack*, Jisoo Kim*, Ki Yong Jung, **DongGeon Lee**, and Heesun Park. Tabular-TX: Theme-Explanation Structure-based Table Summarization via In-Context Learning. In the *Proceedings of the 36th Annual Conference on Human & Cognitive Language Technology (HCLT 2024)*, 2024. (Excellent Paper Award)

[3] Changhun Koo*, Yoonjoo Jung*, and **DongGeon Lee***. Through deep learning-based video processing, Design and implementation of Smart Port Parking Information System. In the *Proceedings of the Annual Conference of Korea Information Processing Society 2021 (ACK 2021)*, 2021.

Under Review

[1] Enhancing Adverse Event Reporting with AI: Using Large Language Models to Detect Inpatient Falls

[2] Decomposing Non-Factoid Question with Retrieval-Augmented Generation

Honors

Gold Prize (Director's Award of the NIKL) Oct 2024
NIKL (National Institute of Korean Language)

- Won the Korean AI Language Proficiency Challenge held by the NIKL.

Excellent Paper Award (Director's Award of the NIKL) Oct 2024
The 36th Annual Conference on Human & Cognitive Language Technology (HCLT 2024)

Scholarship for Outstanding Graduate Students May 2024
POSTECH

Top Engineering Student Award Feb 2024
Inha University

Research Scholarship for Undergraduate Researchers Aug 2023
Inha University