

DongGeon Lee

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

RESEARCH INTERESTS

Data-centric Natural Language Processing, Aligning LLMs/VLMs to build trustworthy AI,
RAG & domain adaptation of LLMs for domain-specific question answering

EDUCATIONS

M.S. student in Artificial Intelligence <i>POSTECH (Pohang University of Science and Technology)</i>	Feb 2024 - Present <i>Pohang, South Korea</i>
--	--

- Advisor: Prof. Hwanjo Yu

B.S. in Information and Communication Engineering <i>Inha University</i>	Mar 2018 - Feb 2024 <i>Incheon, South Korea</i>
--	--

RESEARCH EXPERIENCES

Graduate Research Assistant <i>Data Intelligence Lab, POSTECH</i>	Feb 2024 - Present <i>Pohang, South Korea</i>
---	--

- Research on building a robust RAG (Retrieval-Augmented Generation).
- Research on domain-incremental learning in LLMs (Large Language Models).
- Advisor: Prof. Hwanjo Yu

Undergraduate Research Assistant <i>Data Intelligence Lab, Inha University</i>	Nov 2022 - Nov 2023 <i>Incheon, South Korea</i>
--	--

- Research on domain adaptation of language models in the field of aviation.
- Advisor: Prof. Wonik Choi

Undergraduate Research Assistant <i>Nursing Informatics Lab, Inha University</i>	Jul 2021 - Jun 2023 <i>Incheon, South Korea</i>
--	--

- Research on domain adaptation of language models in the medical field.
- Advisor: Prof. Insook Cho

TECHNICAL SKILLS

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

CONFERENCES (INTERNATIONAL)

- [1] Insook Cho, EunJu Lee, and **DongGeon Lee**. Effects of Language Differences on Inpatient Fall Detection Using Deep Learning. In *Proceedings of the 19th World Congress on Medical and Health Informatics (MedInfo 2024)*, 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, Hyeonsoo Nam, and Yunho Maeng. Question Types Matter: An Analysis of Question-Answering Performance in Retrieval-Augmented Generation Across Diverse Question Types. In *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 *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 *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) <i>NIKL (National Institute of Korean Language)</i> <ul style="list-style-type: none">• Won the Korean AI Language Proficiency Challenge held by the NIKL.	Oct 2024
Excellent Paper Award (Director's Award of the NIKL) <i>The 36th Annual Conference on Human & Cognitive Language Technology (HCLT 2024)</i>	Oct 2024
Scholarship for Outstanding Graduate Students <i>POSTECH</i>	May 2024
Top Engineering Student Award <i>Inha University</i>	Feb 2024
Research Scholarship for Undergraduate Researchers <i>Inha University</i>	Aug 2023