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, Domain adaptation for Large Language Models, Improvement & Evaluation of Retrieval-augmented Language Models

Educations

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

- Advisor: Prof. Hwanjo Yu

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

- Honors : *Research Scholarship for Undergrad. Researcher*, *Top Engineering Student Award*

Research Experiences

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

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

Visiting Researcher Dec 2023 - Jan 2024
HyperEZ Inc. *Seongnam, South Korea*

- Research on the use of RAG (Retrieval-Augmented Generation) in 3D VR platforms.
- Host: Sangseok Youn

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

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

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

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

Conferences

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

[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 KIPS 2021*, 2021. (*: Co-First Author)

Under Review

[1] Question Types Matter: An Analysis of Question-Answering Performance in Retrieval-Augmented Generation Across Diverse Question Types

[2] TXG: Thematic and eXplanatory Generation Pipeline for Table Understanding

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

Technical Skills

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