# 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).

**Research Intern** — Jan 2025 - Feb 2025
*KT Corporation* — *Seoul, South Korea*

- Research on mathematical data synthesis for pre-training LLMs

**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 detecting fall events in clinical notes by 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)

## Publications

[1] Typed-RAG: Type-aware Multi-Aspect Decomposition for Non-Factoid Question Answering
DongGeon Lee*, Ahjeong Park*, Hyeri Lee, Hyeonseo Nam, Yunho Maeng
NAACL'25 SRW | Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (Student Research Workshop)

[2] REFIND: Retrieval-Augmented Factuality Hallucination Detection in Large Language Models
DongGeon Lee, Hwanjo Yu
Preprint'25 | arXiv preprint arXiv:2502.13622

[3] Enhancing Adverse Event Reporting With Clinical Language Models: Inpatient Falls
Insook Cho, Hyunchul Park, Byeong Sun Park, DongGeon Lee
Journal of Advanced Nursing (SCIE; Q1), 2025.02

[4] Theme-Explanation Structure for Table Summarization using Large Language Models: A Case Study on Korean Tabular Data
TaeYoon Kwack*, Jisoo Kim*, Ki Yong Jung, DongGeon Lee, Heesun Park
Preprint'25 | arXiv preprint arXiv:2501.10487

[5] Question Types Matter: An Analysis of Question-Answering Performance in Retrieval -Augmented Generation Across Diverse Question Types
DongGeon Lee*, Ahjeong Park*, Hyeri Lee, Hyeonseo Nam, Yunho Maeng
HCLT'24 (Domestic) | Annual Conference on Human & Cognitive Language Technology

[6] Effects of Language Differences on Inpatient Fall Detection Using Deep Learning
Insook Cho, EunJu Lee, DongGeon Lee
MedInfo'23 | World Congress on Medical and Health Informatics

[7] Bridging the Reporting Gap of Inpatient Falls to Improve Safety Practices Using Deep -Learning-Based Language Models and Multisite Data
DongGeon Lee, EunJu Lee, Insook Cho
CIC'23 | AMIA 2023 Clinical Informatics Conference

## 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*