YONGSIN PARK

yongsin.nlp@gmail.com • https://yongsinp.github.io

SUMMARY

Strong background in natural language processing (NLP), information retrieval (IR), information extraction (IE), and machine learning (ML). Focused on question answering and making information accessible for everyone.

PROFESSIONAL & RESEARCH EXPERIENCE

NLP Research Lab, NHN Diquest | #1 search engine and chatbot provider in Korea — Seoul, South Korea
Senior Researcher, Foundational Technologies Team — *Mar. 2022 - Jun. 2023*
Associate Researcher, Foundational Technologies Team — *Feb. 2019 - Feb. 2022*

- ○ **INFORMATION RETRIEVAL & EXTRACTION**
- · (Project Manager) Developed a content recommendation system based on user behavior, serving 5M users.
- · (Project Manager) Led conversation summarization for the legal domain using deep learning (DL) and reinforcement learning (RL), a system adopted by the Korea Legal Aid Corporation to assist legal consultations.
- · Worked on machine reading comprehension (MRC) for Korean question answering (QA), incorporated into our core chatbot product.
- · Built a recommendation system for course and job opportunity matching used by 35,000 university students.
- ○ **LANGUAGE ANALYSIS**
- · Developed and maintained high-throughput part-of-speech (POS) tagger, named entity recognition (NER) model, and speech act (SA) and sentiment classifier used by all core company products, including a search engine that processes millions of queries per day and high-traffic chatbots.
- · Built a customizable, hybrid NER model using DL and pattern matching. *1st Place, U+ BMT NER*
- · Improved spelling and spacing correction models using statistical post-processing, reducing errors by 47%.
- ○ **OTHERS**
- · Developed a self-service **MLOps** platform for managing diverse ML models, abstracting complex infrastructure to enable non-expert teams to train and deploy solutions on site.
- · Developed data security and NLP tools for personal data extraction and masking, and sentence completion, used by the Korean National Police Agency to ensure compliance and accelerate criminal investigations.

University of Washington / Intelligent Systems Lab, Dong-A University

- ○ **INFORMATION RETRIEVAL & EXTRACTION**
- · Built a lay summarization system for medical texts by fine-tuning Llama 3. *1st Place, BioNLP @ ACL 2025*
- · Created a tool to generate emergency alerts in 15 languages by fine-tuning and constraining Llama 3.
- · Developed a Chrome extension to improve web accessibility for people with visual impairments using Llama 4.
- · Worked on MRC for English QA using language analysis results. *Master's Thesis*
- ○ **LANGUAGE ANALYSIS**
- · Built DL models for POS tagging, NER, dependency parsing (DP), and semantic role labeling (SRL) for a comprehensive and streamlined language analysis. *Bronze Prize, Korean Information Processing System Competition*
- · Built a SA classifier combining hand-crafted rules and a DL model. *Proceedings of the 30th Annual Conference on Human and Cognitive Language Technology*

EDUCATION

Master of Science in Computational Linguistics (NLP) (3.97/4.0) — Seattle, USA
University of Washington — *Sep. 2024 - Present*

Master of Science in Computer Engineering (3.91/4.0) — Busan, South Korea
Dong-A University — *Sep. 2017 - Feb. 2019*

Bachelor of Engineering in Computer Engineering (3.97/4.0) — Busan, South Korea
Dong-A University — *Mar. 2014 - Aug. 2017*

SKILLS

Python • Java • Go • C/C++ • Bash • TensorFlow • PyTorch • scikit-learn • NumPy • Pandas • Parameter Efficient Fine-tuning • Regex • Git/Gitflow • Docker • SQL • Databases (MariaDB, MySQL, Oracle DB)