

SOYOUNG YOON

Ph.D. Student
NLP & Information Retrieval

soyoung97.github.io/profile
soyoung.yoon@snu.ac.kr
Google scholar

Research Interest

My main research goal is to develop NLP systems that are beneficial and informative to people in everyday life. To achieve this, I am focusing on endowing language models with the ability to: (1) Efficiently stay up-to-date with new knowledge and human feedback, (2) Incorporate internal (commonsense) knowledge with external knowledge, and (3) Generate reliable responses grounded on external knowledge base.

Education

Ph.D. in Interdisciplinary Program in Artificial Intelligence <i>Seoul National University (SNU)</i> • Advisor: Prof. Seung-won Hwang , GPA 4.0/4.3	Sep 2023 – present South Korea
M.S. in Kim Jaechul Graduate School of AI <i>Korea Advanced Institute of Science and Technology (KAIST)</i> • Advisor: Prof. Minjoon Seo , GPA 4.1/4.3	Mar 2021 – Aug 2023 South Korea
B.S. in Computer Science, completed courses focused on AI.(Advanced Major) <i>Korea Advanced Institute of Science and Technology (KAIST)</i> • GPA 3.77/4.3 (Cum Laude), Major only 3.92/4.3	Mar 2016 – Feb 2021 South Korea

Publications

(P: preprint, C: conference, J: journal, W: workshop, *: equal contribution)

- [P1] **AcuRank: Uncertainty-Aware Adaptive Computation for Listwise Reranking**
[Soyoung Yoon*](#), [Gyuwan Kim*](#), [Gyu-Hwang Cho](#), [Seung-won Hwang](#)
arxiv preprint
- [C6] **RoToR: Towards More Reliable Responses for Order-Invariant Inputs**
[Soyoung Yoon](#), [Dongha Ahn](#), [Youngwon Lee](#), [Minkyu Jung](#), [HyungJoo Jang](#), [Seung-won Hwang](#)
ACL 2025
- [W1] **Analyzing the Effectiveness of Listwise Reranking with Positional Invariance on Temporal Generalizability**
[Soyoung Yoon*](#), [Jongyoon Kim*](#), [Seung-won Hwang](#)
CLEF 2024 LongEval Working Notes
- [C5] **ListT5: Listwise Reranking with Fusion-in-Decoder Improves Zero-shot Retrieval**
[Soyoung Yoon](#), [Eunbi Choi](#), [Jiyeon Kim](#), [Hyeongu Yun](#), [Yireun Kim](#), [Seung-won Hwang](#)
ACL 2024 (oral, acceptance rate 2.3%)
- [C4] **Exploring the Practicality of Generative Retrieval on Dynamic Corpora**
[Chaeun Kim*](#), [Soyoung Yoon*](#), [Hyunji Lee](#), [Joel Jang](#), [Sohee Yang](#), [Minjoon Seo](#).
EMNLP 2024, also appeared on GenIR@SIGIR 2024
- [C3] **An Integrated Search System for Korea Weather Data**
[Jinkyung Jo](#), [Dayeon Ki](#), [Soyoung Yoon](#), [Minjoon Seo](#)
EMNLP 2023 Industry Track
- [C2] **Towards standardizing Korean Grammatical Error Correction: Datasets and Annotation**
[Soyoung Yoon](#), [Sungjoon Park](#), [Gyuwan Kim](#), [Junhee Cho](#), [Kihyo Park](#), [Gyu Tae Kim](#), [Minjoon Seo](#), [Alice Oh](#).
ACL 2023
- [C1] **SSMix: Saliency-based Span Mixup for Text Classification**
[Soyoung Yoon*](#), [Gyuwan Kim*](#), [Kyumin Park](#)
ACL 2021 (Findings), Also appeared on the 6th Rep4NLP workshop

Work Experience

Channel Corporation AI team <i>Research Intern (Mentor: Minkyu Jung)</i> <ul style="list-style-type: none">• Worked on Table RAG (blog post)• Published a paper at ACL 2025 [C5] about mitigating the lost-in-the middle problem.	Jan – Aug 2024 South Korea
LG AI Research Exaone Lab <i>Research Intern (Mentors: Yireun Kim, Hyeongu Yun)</i> <ul style="list-style-type: none">• Published a paper at ACL 2024 [C4] about listwise reranking.	May – Nov 2023 South Korea
Hyundai AIRS (Now 42dot.ai) <i>Research Intern (Mentor: Yekyung Kim)</i> <ul style="list-style-type: none">• Worked on text-to-SQL semantic parsing.	Dec 2020 – Feb 2021 South Korea
Clova AI, Naver Corp. <i>Research Intern (Mentor: Gyuwan Kim)</i> <ul style="list-style-type: none">• Published a paper at ACL 2021 Findings [C1] on data augmentation for NLP.	July 2020 - Jan 2021 South Korea
KAIST U&I Lab <i>Undergraduate Research Intern (Mentors: Prof. Alice Oh, Sungjoon Park)</i> <ul style="list-style-type: none">• Published a paper at ACL 2023 [C2] on Korean Grammatical Error Correction.	Aug 2019 – July 2020 South Korea
AITRICS (page) <i>SW Engineer Intern (Mentor: Jaeman An)</i> <ul style="list-style-type: none">• Worked on both front-end and back-end, using Django-Rest framework and Vue.js.• Led to a talk at pycon 2019 (presentation)	Jan 2019 – Aug 2019 South Korea

Honors & Awards

Korea National Science & Technology Scholarship <i>Korea Ministry of Science and ICT</i> <ul style="list-style-type: none">• Won about \$500 per semester.	2018-2021
3rd prize, Undergraduate Reserach Project (URP) Excellence Award <i>Korea Advanced Institute of Science and Technology (KAIST)</i> <ul style="list-style-type: none">• Presented the Korean GEC paper [C2].	2020

Skills

Languages: English (Conversational), Korean (Native), Japanese (Introductory)
TOEFL: 108/120 (Nov. 2018), TOEIC: 975/990 (June. 2020), TEPS: 486/600 (Apr. 2023)

Invited Talks & Blog Posts & Community contributions

Table QA with LLMs - huge table parser by pseudocode filtering (blog post) <i>Published a tech insight blog post about my internship project.</i>	July 2024
Huggingface contributor (pull request) <i>Pull request to fix the t5x conversion script merged at huggingface main.</i>	Oct 2023
KENS(KAIST Educational Network System) contributor (pull request) <i>KENS is an official project for KAIST CS341: Introduction to Computer Networks, which is a 4-credit project-oriented course that KAIST CS undergraduates take.</i>	June 2020
Django Query Optimization for Real-time medical AI data processing (presentation) <i>@ Pycon 2019</i> <i>Gave a talk about the internship project done at Aitrics.</i>	Aug 2019

Academic Services

Conference Reviewer: ACL Rolling Review (ARR) Dec 2023, June 2024, Aug 2024, Oct 2024, Dec 2024
European Chapter of the Association for Computational Linguistics (EACL) 2023
Automated Knowledge Base Construction (AKBC) 2022

Projects

Developing an AI-based forecast support solution.	2021
<i>Korea National Institute of Meteorological Sciences (NIMS)</i>	
Effective NLU methods in resource-constrained environments.	2022
<i>Hyundai Motor Group</i>	

References

Seung-won Hwang , Professor at Seoul National University (SNU)	seungwonh@snu.ac.kr
--	--