SUNKYOUNG KIM

2428, E3-1, 291 Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea 34141
sunkyoung@kaist.ac.kr | sunkyoung.github.io

RESEARCH INTERESTS

Large-scale Language Model, Continual Learning, Machine Learning for NLP

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST) Mar 2020 - Present

· M.S. in Computer Science *Daejeon, Korea*
· Advisor : Minsoo Kim
· Thesis title : An Effective Data Selection Method for Pre-training a Korean Language Understanding Model

Ewha Womans University Mar 2015 - Feb 2020

· B.S. in Library & Information Science *(Magna Cum Laude)* *Seoul, Korea*
· B.S. in Multimedia
· Minor : Computer Science & Engineering

Nanyang Technological University June 2019 - July 2019

· Exchange Student *Singapore*

RESEARCH PROJECTS

ITRC Center Student-led Creative Project Mar 2021 - Dec 2021

· Efficient text similarity calculation and deduplication study for building neural network of large-scale text graph
· Participated in data processing for deduplication of large-scale News data and evaluation deduplicated text data on pre-training BERT-based model

KLUE Benchmark Project Dec 2020 - Jul 2021

· Dataset construction of eight Korean natural language understanding downstream tasks
· Participated in dataset construction of KLUE-RE, KLUE-MRC, modeling of KLUE-MRC, and code refactoring for KLUE-MRC finetuning pipeline

PUBILICATIONS

KLUE: Korean Language Understanding Evaluation

· Sungjoon Park, Jihyung Moon, Sungdong Kim, Won Ik Cho, Jiyoon Han, Jangwon Park, Chisung Song, Junseong Kim, Yongsook Song, Taehwan Oh, Joohong Lee, Juhyun Oh, Sungwon Lyu, Younghoon Jeong, Inkwon Lee, Sangwoo Seo, Dongjun Lee, Hyunwoo Kim, Myeonghwa Lee, Seongbo Jang, Seungwon Do, **Sunkyoung Kim**, Kyungtae Lim, Jongwon Lee, Kyumin Park, Jamin Shin, Seonghyun Kim, Lucy Park, Alice Oh, Jung-Woo Ha, and Kyunghyun Cho
· Conference on NeurIPS 2021 Datasets and Benchmarks Track

An efficient and effective document deduplication by using similarity based clustering

· Jungbin Son, **Sunkyoung Kim**, and Min-Soo Kim
· Conference on Korea Computer Congress 2021

TEACHING EXPERIENCE

CS492A Distributed Transactions and Blockchain, KAIST Fall 2020
Teaching Assistant

SCHOLARSHIP

Ewha Future Planning Scholarship 2019
Future Innovation Center, Ewha Womans University

Study Abroad Program Scholarship 2019
International Affair Office, Ewha Womans University

Dean's List 2016, 2017, 2019

Samsung Convergence Software Course Excellence Scholarship 2017
Merit-based Scholarship in Ewha-Samsung Convergence Software Course

Leadership Scholarship 2017
Department of Library and Information Science, Ewha Womans University

Ewha Volunteering Scholarship 2017
Ewha Womans University

Yoo So Young Alumni Scholarship 2017
Department of Library and Information Science, Ewha Womans University

TECHNICAL SKILLS

Programming Languages	Python
Machine Learning Framework	Tensorflow, PyTorch
Tools	Git/GitHub, LaTeX

REFERENCE

Prof.Min-Soo Kim, School of Computing, KAIST, minsoo.k@kaist.ac.kr