Juhwan Choi

gold5230@cau.ac.kr | c-juhwan | c-juhwan.github.io

Research Interest

Deep Learning, Natural Language Processing, Large Language Models, Data-Centric AI

Education

Chung-Ang University — Seoul, Korea
M.S. in Artificial Intelligence — Mar. 2023 - Feb. 2025 (Expected)
GPA: 4.33 / 4.5
Advisor: YoungBin Kim

Chung-Ang University — Seoul, Korea
B.S. in Electrical and Electronics Engineering — Mar. 2017 - Feb. 2023
GPA: 3.94 / 4.5 [Certificate] [Transcript]

Research Experience

Research Assistant — Intelligent Information Processing Lab @ Chung-Ang Univ.
Advisor: YoungBin Kim — Mar. 2023 - Present

Undergraduate Research Intern — Intelligent Information Processing Lab @ Chung-Ang Univ.
Advisor: YoungBin Kim — Apr. 2021 - Feb. 2023

Undergraduate Intern — Big Data & AI Lab @ Chung-Ang Univ.
Advisor: Hohyun Park — Jan. 2021 - Feb. 2021

International Publications

Domain Generalizaton for Dataset Synthesization
Under review.

Interactive Document Summarization System
Under review.

Constructing Sport Game Conversation Dataset
Under review.

Multi-News+: Cost-efficient Dataset Cleansing via LLM-based Data Annotation [Paper]
Juhwan Choi, Jungmin Yun, Kyohoon Jin and YoungBin Kim
arXiv Preprint arxiv:2404.09682

Don't be a Fool: Pooling Strategies in Offensive Language Detection from User-Intended Adversarial Attacks [Paper]
Seunguk Yu, Juhwan Choi and YoungBin Kim
NAACL 2024 Findings

Enhancing Effectiveness and Robustness in a Low-Resource Regime via Decision-Boundary-aware Data Augmentation [Paper]
Kyohoon Jin*, Junho Lee*, Juhwan Choi*, Sangmin Song* and YoungBin Kim
LREC-COLING 2024
Equal contribution as co-first author.

Colorful Cutout: Enhancing Image Data Augmentation with Curriculum Learning [Paper] [Code]
Juhwan Choi and YoungBin Kim
ICLR 2024 Tiny Papers

Adverb Is the Key: Simple Text Data Augmentation with Adverb Deletion [Paper] [Code]
Juhwan Choi and YoungBin Kim
ICLR 2024 Tiny Papers

AutoAugment Is What You Need: Enhancing Rule-based Augmentation Methods in Low-resource Regimes [Paper] [Code] [Poster]
Juhwan Choi, Kyohoon Jin, Junho Lee, Sangmin Song and YoungBin Kim
EACL 2024 Student Research Workshop

GPTs Are Multilingual Annotators for Sequence Generation Tasks [Paper] [Code] [Poster]
Juhwan Choi, Eunju Lee, Kyohoon Jin and YoungBin Kim
EACL 2024 Findings

SoftEDA: Rethinking Rule-Based Data Augmentation with Soft Labels [Paper] [Review] [Code]
Juhwan Choi, Kyohoon Jin, Junho Lee, Sangmin Song and YoungBin Kim
ICLR 2023 Tiny Papers
Accepted as notable work with oral presentation.

Generative Data Augmentation via Wasserstein Autoencoder for Text Classification [Paper] [Code]
Kyohoon Jin, Junho Lee, Juhwan Choi, Soojin Jang and YoungBin Kim
The 13th International Conference on ICT Convergence

Domestic Publications

텍스트 분류를 위한 Wasserstein Autoencoder를 활용한 생성적 데이터 증강
Kyohoon Jin, Junho Lee, Juhwan Choi, Sangmin Song, Soojin Jang and YoungBin Kim
KICS Winter Conference 2024

소프트 라벨을 적용한 규칙 기반 텍스트 데이터 증강 기법 [Paper] [Poster]
Juhwan Choi, Junho Lee, Sangmin Song, Kyohoon Jin and YoungBin Kim
2023 Summer Annual Conference of IEIE

감정 간의 관계를 고려한 지도 대조 학습 기반 감정 인식 [Paper]
Dongje Yoo, Kyunghoon Jeon, Juhwan Choi and YoungBin Kim
2023 Summer Annual Conference of IEIE

KITE: 한국어 고의 오탈자를 활용한 텍스트 데이터 증강 방법론 [Paper]
Seunguk Yu, Juhwan Choi, Heejae Suh, Kyohoon Jin and YoungBin Kim
Proceedings of HCI Korea 2023

낯선 데이터를 활용한 과잉신뢰 완화 텍스트 증강 기법 [Paper] [Presentation]
Junho Lee, Sangmin Song, Juhwan Choi, Juhyoung Park, Kyohoon Jin and YoungBin Kim
Proceedings of HCI Korea 2023

Semantic Preservation and Natural Language Data Augmentation via Variational Autoencoder [Paper]
Yuchul Shin, Kyohoon Jin, Juhwan Choi, Junho Lee, Soojin Jang and YoungBin Kim
TechArt

Variational Autoencoder기반 의미 보존 자연어 데이터 증강 기법 [Paper] [Presentation]
Juhwan Choi, Junho Lee, Kyohoon Jin, Yehoon Jang, Soojin Jang and YoungBin Kim
2022 Summer Annual Conference of IEIE

Research Projects

Multimodal Dataset Generation Nov. 2023 - Present

Domain Generalization through Dataset Generation Nov. 2023 - Present
Submitted a paper - under review

Multilingual Data Annotation using LLMs Apr. 2023 - Jan. 2024

Submitted a paper and accepted to EACL 2024 Findings.

Text Data Augmentation with Soft Label Feb. 2023 - May. 2023
Submitted a paper and accepted to ICLR 2023 Tiny Papers.

Interactive News Summarization Jan. 2023 - Present
Submitted a paper at demo track.

Decision Boundary-Based Data Augmentation Aug. 2022 - Feb. 2024
Submitted a paper and accepted to LREC-COLING 2024.

Deep Learning Seminar Feb. 2022 - Mar. 2022
Based on Stanford CS224n and CS231n

Text Data Augmentation using Wasserstein Autoencoder Aug. 2021 - Jan. 2022
Related work research, Implemented text classification model.

Implementing Image Captioning Model using Transformers Apr. 2021 - Jun. 2021

Korean Hate Speech Detection Model Apr. 2021 - May. 2021
Used pre-trained KcELECTRA. Ranked 11/88 (in Dec. 2022) at [Kaggle competition]

Deep Learning Seminar Jan. 2021 - Feb. 2021
Based on CMU 11-785

Honors & Awards

Travel Grant @ ICLR 2024

Skills

Programming Languages **Familiar:** Python | C | C++
Experienced: Matlab
Libraries & Tools Pytorch | HuggingFace | Numpy | Pandas | Git | CLI
Language Proficiency **Native:** Korean
Fluent: English [TOEIC 955] [TEPS 409]

Services

Reviewer ACL Rolling Review (From 2023 August)
EACL 2024
NAACL 2024
ACL 2024
ICLR 2024 Tiny Papers

Relevant Coursework

Graduate Level

Artificial Intelligence Deep Unsupervised Learning | Deep Generative Model | Graph Neural Networks | Advanced Natural Language Processing | Artificial Intelligence Literacy | Ethics and Fairness in Artificial Intelligence

Undergraduate Level

Computer Science Algorithms | Artificial Intelligence | Basic Computer Programming | Computer Architecture | Data Structure | Digital Image Processing | Internet of Intelligent Things | Logic Circuit | Microcomputer Systems | Mobile Network | Numerical Analysis | Object-Oriented Programming | Operating Systems
Mathematics Calculus 1 | Calculus 2 | Electric Mathematics | Engineering Mathematics | Linear Algebra | Random Variables