Shinhyeok Oh

NLP RESEARCH SCIENTIST - ALTERNATIVE MILITARY SERVICE (JUL 2021 - JUL 2024)

Seoul, Republic of Korea

☐ (+82) 10-3125-7712 | ✉ achromatically@naver.com | 🏠 shinhyeokoh.github.io | 🔗 shinhyeok-oh-5082b51a8 | 🎓 Google Scholar

Work Experience

Riiid
Seoul, Republic of Korea

NLP RESEARCH SCIENTIST (ALTERNATIVE MILITARY SERVICE)
Dec 2022 - Present

- Large Language Model
  - research on cross-lingual LLM (English, Korean)
  - single-node & multi-node training using DeepSpeed and FSDP (- 70b)
- Automated Essay Assessment
  - research on alleviating performance degradation in scoring long-text
  - designing self-consistency methods (e.g. prompt decomposition) for in-context learning pipeline
  - successful execution of PoC related to AP World History and Bar Exam
- Question Generation & Evaluation
  - research on robust QG evaluation approaches for Riiid's AI product (Quizium)
  - published in ACL 2023 (1st author; findings)
- Japanese Question Auto-tagging
- TOEIC Writing Scoring

Netmarble
Seoul, Republic of Korea

NLP RESEARCHER & ENGINEER (ALTERNATIVE MILITARY SERVICE)
Apr 2021 - Dec 2022

- Game NMT Service
  - providing machine translation service for the global release of new games
  - maintenance on Magellan Machine Translation (MaMT) system (springboot + docker + jenkins)
  - research on glossary modeling
- Automatic Post-editing
  - research on automatic post-editing task (en-de)
  - 1st place, automatic post-editing task @WMT, EMNLP 2021
- Text-to-Speech Synthesis
  - research on multi-modal modeling using dependency parsing feature
  - published in AAAI 2023 (1st author; oral)
- Chatbot
  - research on persona chatbot, generative chatbot w/ passage retrieval

Kakao Enterprise
Seongnam, Republic of Korea

NLP RESEARCHER (RESEARCH INTERNSHIP)
Sep 2020 - Feb 2021

- Aspect-based Sentiment Analysis
  - research on aspect-based sentiment analysis using self-supervised and multi-task learning
  - published in ACL 2021 (1st author; oral)

Selected Publications from Industry

Evaluation of Question Generation Needs More References

Shinhyeok Oh*, Hyojun Go*, Hyeongdon Moon, Yunsung Lee, Myeongho Jeong, Hyun Seung Lee and Seungtaek Choi (**equal contribution*)

ACL 2023 (Findings, short) - @Riiid

RWEN-TTS: Relation-aware Word Encoding Network for Natural Text-to-Speech Synthesis

Shinhyeok Oh*, HyeongRae Noh*, Yoonseok Hong, Insoo Oh

AAAI 2023 (Oral Presentation) - @Netmarble

Netmarble AI Center's WMT21 Automatic Post-Editing Shared Task Submission

Shinhyeok Oh*, Sion Jang*, Hu Xu*, Shounan An, Insoo Oh

WMT @EMNLP 2021 (1st place) - @Netmarble

Deep Context- and Relation-Aware Learning for Aspect-based Sentiment Analysis

Shinhyeok Oh*, Dongyub Lee*, Taesun Whang, IlNam Park, Gaeun Seo, EungGyun Kim, Harksoo Kim

ACL 2021 (Oral Presentation, short) - @Kakao Enterprise

Honors & Awards

OCT 2023 **1st Place**, Open LLM Leaderboard (as of Oct 4, 2023) @Huggingface
Nov 2021 **1st Place**, Automatic Post-Editing shared task @WMT, EMNLP 2021
Oct 2020 **Special Prize**, Korean Language Information Processing Competition - Chatbot @HCLT 2020
Oct 2020 **Special Prize**, Korean Language Information Processing Competition - Sentiment Analysis @HCLT 2020
Oct 2019 **Bronze Prize**, Korean Language Information Processing Competition - Dependency Parsing @HCLT 2019
Sep 2019 **Gold Prize**, Kangwon SW Festival - @National Program of Excellence in Software
Oct 2018 **Grand Prize**, Korean Language Information Processing Competition - Dependency Parsing @HCLT 2019
Jul 2016 **1st Place**, Department of Computer and Communications Engineering (Rank: 1/97), Spring Semester 2016

Teaching Experience

2023 **Instructor**, Special Lecture for Junior (Guidance on career path) @Kangwon National University *Republic of Korea*
2023 **Instructor**, Special Lecture for Freshman (AI Foundation Models) @Kangwon National University *Republic of Korea*
2022 **Instructor**, Special Lecture for Junior (Career in Artificial Intelligence) @Kangwon National University *Republic of Korea*
2020 **Teaching Assistant**, Natural Language Processing @LG Uplus *Republic of Korea*
2018-2019 **Teaching Assistant**, Java Programming, C Programming, App Programing @Kangwon National University *Republic of Korea*
2017 **Instructor**, Java Programming (Supplementary classes supported by Kangwon National University) *Republic of Korea*

Academic Activities

- Reviewer of ACL, NAACL, EACL, and ARR
- Reviewer of IEEE Transactions on Audio, Speech and Language Processing

Education

Kangwon National University *Chuncheon, Republic of Korea*
M.S. in Department of Computer and Communications Engineering, GPA 4.31 / 4.5 *Sep 2019 - Feb 2021*

Kangwon National University *Chuncheon, Republic of Korea*
B.S. in Department of Computer and Communications Engineering, GPA 3.86 / 4.5 *Mar 2016 - Feb 2021*

Experience as a Student

NLP Lab (Konkuk University) *Seoul, Republic of Korea*
Student (Advisor: Prof. Harksoo Kim) *Mar 2020 - Sep 2020*
- research on persona-based chatbot
- development of Korean language model

NLP Lab (Kangwon National University) *Chuncheon, Republic of Korea*
Student (Advisor: Prof. Harksoo Kim) *Aug 2018 - Mar 2020*
- research on persona-based chatbot
- development of web, demo, annotation tool, gpubot, etc...

Self-employed (using freelancer platform) *Republic of Korea*
Developer *2018-2019*
- development of Web, Android, Application Software
  - C# Form-based GUI Softwares
  - single-page websites
  - Android applications
  - crawling softwares
- 30+ Projects (revenue exceeding 20 million won)

Skills

- Languages: Python, Java, C, C#, HTML/CSS/JS, Android
- Deep Learning Libs: PyTorch, Tensorflow
- Etc.: Springboot, Gnuboard, Google Cloud Platform, Deepspeed, ...

Certificate

Certificate of Engineer Information Processing   *Republic of Korea*

HUMAN RESOURCES DEVELOPMENT SERVICE OF KOREA   *2019*

Other Publications

INTERNATIONAL PUBLICATIONS

Multi-Architecture Multi-Expert Diffusion Models

Yunsung Lee*, Jin-Young Kim*, Hyojun Go*, Myeongho Jeong, Shinhyeok Oh, Seungtaek Choi   (**equal contribution*)

AAAI 2024 (accepted)

Addressing Negative Transfer in Diffusion Models

Hyojun Go*, JinYoung Kim*, Yunsung Lee*, Seunghyun Lee, Shinhyeok Oh, Hyeongdon Moon, Seungtaek Choi

NeurIPS 2023

Cross Encoding As Augmentation: Towards Effective Educational Text Classification

Hyun Seung Lee*, Seungtaek Choi*, Yunsung Lee, Hyeongdon Moon, Shinhyeok Oh, Myeongho Jeong, Hyojun Go, Christian Wallraven

ACL 2023 (Findings, short)

Improving Document-Level Sentiment Classification Using Importance of Sentences

Gihyeon Choi, Shinhyeok Oh, Harksoo Kim

Entropy, 2020 (SCIE)

Multi-Turn Chatbot Based on Query-Context Attentions and Dual Wasserstein Generative Adversarial Networks

Jintae Kim, Shinhyeok Oh, Oh-Woog Kwon, Harksoo Kim

Appl. Sci., 2019 (SCIE)

DOMESTIC PUBLICATIONS

Method of Reflecting Various Personas in a Chatbot

Shinhyeok Oh, Seokwon Jung, Harksoo Kim

Journal of KIISE, 2021 (KCI)

Movie Revies Sentiment Analysis Considering the Order in which Sentiment Words Appear

HongJin Kim, Damrin Kim, Boeun Kim, Shinhyeok Oh, Harksoo Kim

HCLT 2020

Korean Named Entity Recognition Using ELECTRA and Label Attention Network

HongJin Kim, Shinhyeok Oh, Harksoo Kim

HCLT 2020

Korean Generative Chatbot using Topic Embedding

Shinhyeok Oh, Harksoo Kim

HCLT 2020

Personal Characteristics Classifier for Persona Chatbot Research

Shinhyeok Oh, Harksoo Kim, Jeong-Eom Lee, Seona Kim, Youngmin Park, Myungho Noh

KSC 2019

Personalized Multi-Turn Chatbot Based on Dual WGAN

Shinhyeok Oh, Jintae Kim, Harksoo Kim, Jeong-Eom Lee, Seona Kim, Youngmin Park, Myungho Noh

HCLT 2019

Multi-head Attention and Pointer Network Based Syllables Dependency Parser

HongJin Kim, **Shinhyeok Oh**, *Damrin Kim, Boeun Kim, Harksoo Kim*

HCLT 2019

Persona Reflection Method for Generative Chatbot System

Shinhyeok Oh, *Jintae Kim, Youngmin Park, Seona Kim, Jeong-Eom Lee, Harksoo Kim*

KCC 2019

Korean Dependency Parsing Using ELMo and Multi-head Attention

Seongsik Park, **Shinhyeok Oh**, *Hongjin Kim, Sihyung Kim, Harksoo Kim*

HCLT 2018

Korean Dependency Parsing using Multi-head Attention and Pointer Network

Seongsik Park, **Shinhyeok Oh**, *Hongjin Kim, Harksoo Kim*

HCLT 2018