Dongju Park

• E-mail • LinkedIn • Facebook • Google Scholar • GitHub

WORK EXPERIENCE

NAVER Corporation

- Agent Reasoning & Planning Team
- AI Research Engineer — Feb 2020 – Present
  - Natural Language Processing
  - Large-Scale Language Model
  - Vision-Language Model
  - Semi-Supervised Learning

EDUCATION

Gwangju Institute of Science and Technology (GIST)

- M.S in Electrical Engineering and Computer Science — Mar 2018 – Feb 2020
- Meta-Evolutionary Machine Intelligence Laboratory
  - Focus: Natural Language Processing, Deep Learning, Machine Learning
  - Adviser: Prof. Chang Wook Ahn

Chonnam National University

- B.S. in Industrial Engineering — Mar 2012 – Feb 2018

RESEARCH EXPERIENCE

Meta-Evolutionary Machine Intelligence Laboratory, GIST

- Research Intern — Sep 2017 – Feb 2018

PUBLICATIONS

INTERNATIONAL CONFERENCES

[1] D. Park, S. Ka, KM. Yoo, G. Lee and J. Kang, "HyperT5: Towards Compute-Efficient Korean Language Modeling," *Industry Track of Annual Meeting of the Association for Computational Linguistics: ACL 2023*

[2] KM. Yoo, D. Park, J. Kang, S. Lee and W. Park, "GPT3Mix: Leveraging Large-scale Language Models for Text Augmentation," *Findings of the Association for Computational Linguistics: EMNLP 2021*

[3] B. Kim, HS. Kim, S. Lee, G. Lee, D. Kwak, JD. Hyeon, S. Park, S. Kim, S. Kim, D. Seo, H. Lee, M. Jeong, S. Lee, M. Kim, SH. Ko, S. Kim, T. Park, J. Kim, S. Kang, N. Ryu, KM. Yoo, M. Chang, S. Suh, S. In, J. Park, K. Kim, H. Kim, J. Jeong, YG. Yeo, D. Ham, D. Park, MY. Lee, J. Kang, I. Kang, J. Ha, W. Park, and N. Sung, "What Changes Can Large-scale Language Models Bring? Intensive Study on HyperCLOVA: Billions-scale Korean Generative Pretrained Transformers," *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: EMNLP 2021*

[4] C. Kim, D. Park and HN. Lee, "Convolutional neural networks for the reconstruction of spectra in compressive sensing spectrometers," *SPIE Photonics West 2019*

[5] D. Park and CW. Ahn, "LSTM Encoder-Decoder with Adversarial Network for Text Generation from Keyword," *The 13th International Conference on Bio-inspired Computing: Theories and Applications (BIC-TA 2018)*

INTERNATIONAL JOURNALS

[1] C. Kim, D. Park and HN. Lee, "Compressive Sensing Spectroscopy Using a Residual Convolutional Neural Network," *Sensors*, vol. 20, no. 3: 594, 2020.

[2] D. Park and CW. Ahn, "Self-Supervised Contextual Data Augmentation for Natural Language Processing," *Symmetry*, vol. 11, no. 11: 1393, 2019.

ARXIV

[1] HyperCLOVA X Team, "HyperCLOVA X Technical Report," *arXiv*, 2024.

[2] C. Kim, D. Park and HN. Lee, "Deep learning-based single-shot computational spectrometer using multilayer thin films," *arXiv*, 2022.

DOMESTIC CONFERENCES

[1] D. Park and CW. Ahn, “Named Entity Recognition using Bidirectional LSTM-CRF Combining Named Entity Ratio Dictionary,” *Korea Computer Congress*, 2019.

[2] D. Park and CW. Ahn, “Classifying Documents with Self-Attention Network Built on Input-Keyword Combination,” *Spring Conference Of Korean Institute of Smart Media*, 2019.

[3] D. Park and CW. Ahn, “Sentence Generation from Keyword using Generative Adversarial Networks,” *Korea Computer Congress*, 2018.

DOMESTIC JOURNALS

[1] D. Park, BW. Kim, YS. Jeong, and CW. Ahn, “Deep Neural Network Based Prediction of Daily Spectators for Korean Baseball League : Focused on Gwangju-KIA Champions Field,” *Smart Media Journal*, vol. 7, no. 1, pp. 16–23, Mar 2018.)

PROJECTS

The Development of Harmonics-based Sound Design in view of Driver’s Preference and Driving Condition

- Hyundai Motors — May 2019 – Dec 2019
  - Sound design for driving conditions using deep learning
  - Analysis of personal preference using natural language processing

Distributed Deep Reinforcement Learning for Real-world Problem

- Gwangju Institute of Science and Technology — Mar 2019 – Dec 2019
  - RNN and LSTM model design and hyperparameter tuning for time series data

Co-evolutionary Interaction based Emergent Art Creation System with Multiobjective Aesthetic Evaluation

- National Research Foundation of Korea — Mar 2019 – Dec 2019
  - Implementation of Generative adversarial networks models for comparison with evolutionary algorithms

Evolutionary Neural Network for Object Detection in a Wide Range of Distance for Autonomous Vehicles

- National Research Foundation of Korea — Jul 2018 – Feb 2019
  - CNN and LSTM model design and hyperparameter tuning for time series data

Evolutionary Machine Learning based Emotional Contents Generation

- Gwangju Institute of Science and Technology — Aug 2018 – Dec 2018
  - Deep learning based methodology baseline implementation by implementing various GAN and LSTM based models

AWARDS & SCHOLARSHIPS

3th place of Ressearch Track, N-innovation 2024 @ NAVER

- Open your eyes! CLOVA X — 2024

3th place of Ressearch Track, N-innovation 2022 @ NAVER

- Hyperscale AI-based No Code AI Platform ’CLOVA Studio’ — 2022

CLOVA Wow Project 2021

- HyperCLOVA — 2021

Clova Wow Project 2020

- AiCALL — 2020

1st place, Haafor Challenge 2020 @ HAAFOR

- Finding the Chronological Order of Articles — 2020

4th place, Commercial Online Game Data Analysis Competition @ GIST

- Design for Online Game Churn Prediction Model for considering residual value using the Commercial Online Game Data — 2020

1st place, Naver NLP Challenge 2018 @ NAVER

- Named Entity Recognition Task — 2018

PROFESSIONAL AFFILIATIONS & ACTIVITIES

Tech Talk @ DEVIEW 2023

- Seq2Seq HyperCLOVA is Mostly What You Need: Understanding, Generation, and Efficiency 2023

Tech Talk @ DEVIEW 2021 and MODUCON

- There is no useless data in the world: semi-supervised learning with HyperCLOVA 2021

NVIDIA Deep Learning Institute Instructor

- Fundamentals of Deep Learning for Natural Language Processing 2019

TEACHING EXPERIENCE

NVIDIA Deep Learning Institute @ GIST

- Instructor, Fundamentals of Deep Learning for Natural Language Processing 2020

Software Practical Use and Coding @ GIST

- Teaching Assistant, Data Crwaling and Deep Learning 2019

NVIDIA Deep Learning Institute @ NVIDIA DLI

- Teaching Assistant, Deep Learning Fundamentals for Multi-GPU 2019

Machine Learning and Deep Learning @ KEPCO KDN

- Teaching Assistant, Machine Learning, Deep Learning and Tensorflow 2018

Research and Education @ Jeonnam Science High School

- Teaching Assistant, Creative Font Generation System using Deep Learning 2018

LANGUAGES

- Korean: Native language
- English: Intermediate

SKILLS

- **Python**, Scala, C++, JAVA

[*CV compiled on 2025-07-07*]